

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 15, 2004, 16:35:53 ; Search time 3682 Seconds  
(without alignments)

1671.567 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 778

Sequence: 1 MGAPLPPAWQPFLLKDHRS.....EFETAKKVRRAIEQLAAMD 142

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delcsp 6.0 , Delcxt 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cpn2.1/USPTO.spool/US09690825/runat.11082004.140951.13531/app.query.fasta\_1.327  
-DB=GenEmbl -OPWT=fastap -SUFFIX=rge -VINMATCH=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09690825 @CCN\_1.1.3731 @runat.11082004.140951.13531 -NCPUS=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_hgt\_hum.\*  
31: em\_hgt\_inv.\*  
32: em\_hgt\_other.\*  
33: em\_hgt\_mus.\*  
34: em\_hgt\_pin.\*  
35: em\_hgt\_rod.\*  
36: em\_hgt\_man.\*  
37: em\_hgt\_vrt.\*  
38: em\_sy.\*  
39: em\_hgtgo\_hum.\*  
40: em\_hgtgo\_mus.\*  
41: em\_hgtgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	778	100.0	426	6	AR184473 Sequence
2	778	100.0	426	6	BD167854 Survivin
3	778	100.0	426	6	BD185366 Survivin
4	778	100.0	1619	6	AR093347 Sequence
5	778	100.0	1619	6	BD273550 Antisense
6	778	100.0	1619	6	AR181635 Sequence
7	778	100.0	1619	6	AX775129 Sequence
8	778	100.0	1619	6	AX779941 Sequence
9	778	100.0	1629	9	AF077350 Homo sapi
10	778	100.0	1643	9	BC034148 Homo sapi
11	778	100.0	1653	9	BC008718 Homo sapi
12	775	99.6	1630	4	AB095108 Canis fam
13	756.5	97.2	600	9	AB028869 Homo sapi
14	748	96.1	1165	6	AR097642 Sequence
15	748	96.1	1165	6	AR154245 Sequence
16	721	92.7	794	6	AF270355 Sequence
17	718	92.3	794	6	AF195781 Sus scrofa
18	663.5	85.3	463	6	AX778483 Sequence
19	660	84.8	903	10	AF276775 Rattus no
20	660	84.8	924	10	AF077349 Mus muscu
21	660	84.8	955	6	AR181548 Sequence
22	660	84.8	955	10	AB013819 Mus muscu
23	642	82.5	399	6	AX886452 Sequence
24	642	82.5	399	6	BD026062 Sequence
25	618	79.4	1011	9	HUMMNASEQ
26	556	71.5	3352	10	BC004702 Mus muscu
27	531.5	68.3	1539	9	BC000784 Homo sapi
28	486	62.5	503	5	AF322051 Gallus ga
29	459.5	59.1	555	5	AF377323 Gallus ga
30	456	58.6	772	5	AY174765 Xenopus l
31	389	50.0	701	5	AY057057 Danio rer
32	389	50.0	1239	5	BC056739 Danio rer
33	382	49.1	14796	6	AR157507 Sequence
34	382	49.1	14796	6	AR181541 Sequence
35	382	49.1	14796	6	AX333491 Sequence
36	382	49.1	14796	6	AX334154 Sequence
37	382	49.1	14796	6	AX336483 Sequence
38	382	49.1	14796	6	AX410775 Sequence
39	382	49.1	14796	6	BD192864 Survivin
40	382	49.1	14796	9	RSU75285 Homo sapien
41	382	49.1	209751	9	AC087645 Homo sapi
42	363	46.7	489	5	AY057058 Danio rer
43	363	46.7	154840	2	AC032035 Homo sapi
44	361.5	46.5	228077	2	AC130552 Rattus no
45	361.5	46.5	240200	2	AC122075 Rattus no

#### ALIGNMENTS

RESULT 1

AR184473  
 LOCUS AR184473 426 bp DNA linear PAT 20-APR-2002  
 DEFINITION Sequence 2 from patent US 6346389.  
 ACCESSION AR184473  
 VERSION AR184473.1 GI:20230438  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE  
 1. (bases 1 to 426)  
 AUTHORS Altieri, D.C.  
 TITLE Method for selectively modulating the interactions between survivin and tubulin  
 JOURNAL Patent: US 6346389-A 2 12-FEB-2002;  
 FEATURES  
 source Location/Qualifiers  
 1..426  
 /organism="unknown"  
 /mol\_type="unassigned DNA"  
 ORIGIN  
 Alignment Scores: 3.01e-77 Length: 426  
 Pred. No.: 778.00 Matches: 142  
 Score: 778.00  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-690-825-34 (1-142) x AR184473 (1-426)  
 QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 1 ATGGGTGCGCCGACGTTGGCCCTGCTGGCAGCCCTTCTCAAGGACCAACCGCATCTCT 60  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 61 ACATTCAGAACTGGCCCTTCTTGAGGGCTGGCCCTGACCCGAGCGGATGGCCGAG 120  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
 Db 121 GCTGGCTTCATCCACTGCCCTGCTGAGACGAGCCAGACTTGGCCCTGCTTCTCTG 180  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80  
 Db 181 TTCAGGAGCTGGAGGCTGGAGCCAGATGACGACCCCATAGAGAACCAATAAAGCAT 240  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 241 TCGTCCGGTTGGCGCTTCTTCTGTCAGAGCAGTTTGAAGAAATTACCCCTTGTGAA 300  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 Db 301 TTTTGAACCTGGACAGAGAAAGCCAGCAAAATTGCAAGGAAACCAACAATAAG 360  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaA 140  
 Db 361 AAGAAAGAAATTGAGGAACTGGCAAGAAAGTGGCGGCTGCCATCGAGCAGCTGGCTGCC 420  
 QY 141 MetAsp 142  
 Db 421 ATGGAT 426  
 RESULT 2  
 BD167854  
 LOCUS BD167854 426 bp DNA linear PAT 17-JAN-2003  
 DEFINITION Survivin-like polypeptide and its DNA.  
 ACCESSION BD167854  
 VERSION BD167854.1 GI:27873666  
 KEYWORDS WO 0233071-A/6.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 426)  
 AUTHORS Tanaka, H. and Kaieda, I.  
 TITLE Survivin-like polypeptide and its DNA  
 JOURNAL Patent: WO 0233071-A 6 25-APR-2002;  
 COMMENT TAKEDA CHEMICAL INDUSTRIES LTD, HIROSHI TANAKA, ISAO KAIEDA  
 OS Homo sapiens (human)  
 PN WO 0233071-A/6  
 PD 25-APR-2002  
 PF 18-OCT-2001 WO 2001JP009071  
 PR 17-OCT-2000 JP 00P 316721, 20-DEC-2000 JP 00P 386809 PI  
 HIROSHI TANAKA, ISAO KAIEDA  
 PC C12N15/09, C12N15/12, C07K14/47, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
 10, C12P21/02, C07K16/18, C12Q1/02, C12Q1/68, G01N33/15, G01N33/50, PC  
 A61K31/711,  
 PC A61K38/17, A61K39/395, A61K48/00, A61P35/00, A61P43/00 CC  
 Survivin-like polypeptide and its DNA  
 FH Key Location/Qualifiers  
 FT source 1..426  
 /organism="Homo sapiens (human)"  
 FEATURES  
 source Location/Qualifiers  
 1..426  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Alignment Scores: 3.01e-77 Length: 426  
 Pred. No.: 778.00 Matches: 142  
 Score: 778.00  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-690-825-34 (1-142) x BD167854 (1-426)  
 QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 1 ATGGGTGCGCCGACGTTGGCCCTGCTGGCAGCCCTTCTCAAGGACCAACCGCATCTCT 60  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 61 ACATTCAGAACTGGCCCTTCTTGAGGGCTGGCCCTGACCCGAGCGGATGGCCGAG 120  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
 Db 121 GCTGGCTTCATCCACTGCCCTGCTGAGACGAGCCAGACTTGGCCCTGCTTCTCTG 180  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80  
 Db 181 TTCAGGAGCTGGAGGCTGGAGCCAGATGACGACCCCATAGAGAACCAACAATAAAGCAT 240  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 241 TCGTCCGGTTGGCGCTTCTTCTGTCAGAGCAGTTTGAAGAAATTACCCCTTGTGAA 300  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 Db 301 TTTTGAACCTGGACAGAGAAAGCCAGCAAAATTGCAAGGAAACCAACAATAAG 360  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaA 140  
 Db 361 AAGAAAGAAATTGAGGAACTGGCAAGAAAGTGGCGGCTGCCATCGAGCAGCTGGCTGCC 420  
 QY 141 MetAsp 142  
 Db 421 ATGGAT 426  
 RESULT 3  
 BD185366  
 LOCUS BD185366 426 bp DNA linear PAT 17-JUN-2003



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 14:14:00 ; Search time 16 seconds  
(without alignments)  
853.700 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 778  
Sequence: 1 MGAPLPPAWQFLKDHRS.....EPEETAKVRRRAIEQLAAMD 142

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*

- 1: pir1:\*
- 2: pir2:\*
- 3: pir3:\*
- 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	163	21.0	997	T43523	cutl7 protein - fi
2	160	20.6	4845	T31067	BIR repeat contain
3	157	20.2	308	T37474	apoptosis inhibito
4	155	19.9	1232	A55478	neuronal apoptosis
5	154	19.8	208	T03183	probable apoptosis
6	145.5	18.7	155	T37471	apoptosis inhibito
7	142	18.3	1447	T42628	neuronal apoptosis
8	136.5	17.5	275	A45679	inhibitor-of-apopt
9	136.5	17.5	298	JC7568	kidney inhibitor o
10	133.5	17.2	268	T10304	inhibitor of apopt
11	133.5	17.2	268	A53989	apoptosis-inhibiti
12	131	16.8	618	S68450	apoptosis inhibito
13	130	16.7	155	T30489	apoptosis inhibito
14	128.5	16.5	358	JC5964	apoptosis inhibito
15	125.5	16.1	497	S63544	apoptosis inhibito
16	121	15.6	604	S68449	apoptosis inhibito
17	118	15.2	496	T28409	apoptosis inhibito
18	118	15.2	497	S69545	apoptosis inhibito
19	106.5	13.7	150	T28409	ORF MSV248 probabl
20	106.5	13.7	954	S57108	hypothetical prote
21	99.5	12.8	275	T10310	apoptosis-inhibiti
22	96	11.1	329	T28403	ORF MSV242 probabl
23	85	10.9	329	T47419	hypothetical prote
24	83.5	10.7	385	T19201	hypothetical prote
25	83	10.7	434	A43252	probable transcrip
26	82	10.5	135	A38609	lectin, galactose-
27	80.5	10.3	1096	T48512	hypothetical prote
28	80	10.3	355	T22146	hypothetical prote
29	80	10.3	1390	S51364	sperm tail-specifi

ALIGNMENTS

RESULT 1

T43523

cutl7 protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 21-Jan-2000 #sequence-revision 21-Jan-2000 #text\_change 02-Jun-2000

C:Accession: T43523; T41649; T41700

R:Morishita, J.; Matsusaka, T.; Yanagida, M.

submitted to the EMBL Data Library, August 1999

A:Description: Fission yeast cutl7 is required for chromosome segregation.

A:Reference number: 222536

A:Accession: T43523

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-997 <MOR>

A:Cross-references: EMBL:AB031034; PIDN:BAAG3415.1

R:Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, August 1998

A:Reference number: 222007

A:Accession: T41649

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-997 <HAR>

A:Cross-references: EMBL:AL031323; PIDN:CAA20434.1; GSPDB:GN00068; SPDB:SPCC962.02c

R:Medler, H.; Duesternoeft, A.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, October 1999

A:Reference number: 222010

A:Accession: T41700

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 932-937 <WED>

A:Cross-references: EMBL:AL121859; PIDN:CAB58376.1; GSPDB:GN00068; SPDB:SPCP31B10.10c

C:Genetics:

A:Gene: cutl7; SPCC962.02c; SPDB:SPCP31B10.10c

A:Map position: 3L

A:Introns: 43/3

Query Match 21.0%; Score 163; DB 2; Length 997;

Best Local Similarity 41.6%; Pred.No. 3.1e-06;

Matches 32; Conservative 9; Mismatches 32; Indels 4; Gaps 2;

QY 18 RISTF--KNWPFLE--GCATPRMAMAGTHCPTENEPDLAOCFFCKLEGWEPDDDP 73

120 REQTFVQKWFYTNRPDYHCEPSVMAASGVYNYNTADAKAAHCLYCDINLHDEWEPDDP 179

QY 74 IEHKKHSSGCAFLSVK 90

180 YTEHKRRRADCVPFTWK 196

Db

Db

RESULT 2

T31067

BIR repeat containing ubiquitin-conjugating enzyme BRUCE - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 21-Jul-2000  
 C:Accession: T31067  
 R:Hauser, H.P.; Bardeff, M.; Pyrowolakis, G.; Jentsch, S.  
 J. Cell Biol. 141, 1415-1422, 1998  
 A:Title: A giant ubiquitin-conjugating enzyme related to IAP apoptosis inhibitors.  
 A:Reference number: Z20963; MUID:98292517; PMID:9628897  
 A:Accession: T31067  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-4845 <HAU>  
 A:Cross-references: EMBL:Y17267; NID:g3319989; PIDN:CAA76720.1; PID:g3319990  
 A:Note: localized to the Golgi compartment and the vesicular system  
 C:Keywords: Membrane-associated protein

Query Match 20.6%; Score 160; DB 2; Length 4845;  
 Best Local Similarity 42.3%; Pred. No. 2.8e-05;  
 Matches 30; Conservative 10; Mismatches 31; Indels 0; Gaps 0;  
 QY 17 HRISTFKNWPFLEGCACCTPERMAEAGFIHCPTENEPDLAOCFFCFKELEGWEPDDDDP 76  
 DB 263 NRRETFSTWPHGVRYMAQDPMAQAGFYHQFASGGDDRAMCFTCSVCLVCWEPTDEPWE 322  
 QY 77 HKHSHSGCAFL 87  
 DB 323 HERHSPNCPV 333

RESULT 3  
 T37474  
 C:Species: Caenorhabditis elegans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
 C:Accession: T37474; T20098  
 R:Uren, A.G.  
 submitted to the EMBL Data Library, September 1996  
 A:Reference number: Z21711  
 A:Accession: T37474  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-308 <URE>  
 A:Cross-references: EMBL:U72208; PIDN:AAD00182.1  
 R:Percy, C.  
 submitted to the EMBL Data Library, July 1996  
 A:Reference number: Z19223  
 A:Accession: T20098  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-308 <WIL>  
 A:Cross-references: EMBL:Z77654; PIDN:CAB01130.1; GSPDB:GN00023; CESP:C50B8.2  
 A:Experimental source: clone C50B8  
 C:Genetics:  
 A:Gene: C50B8.2  
 A:Map position: 5  
 A:Introns: 48/3; 143/1; 191/3

Query Match 20.2%; Score 157; DB 2; Length 308;  
 Best Local Similarity 30.1%; Pred. No. 3.1e-06;  
 Matches 37; Conservative 27; Mismatches 49; Indels 10; Gaps 3;  
 QY 16 DHRISTFKNWPF--LEGACCTPERMAEAGFIHCPTENEPDLAOCFFCFKELEGWEPDDDDP 73  
 DB 168 DHRATFQNFIFDKRNVKYTSKLLAKAGWFSIANKDKTSKAPCPCLVELD--FDSDDP 226

QY 74 IEHKKHSSGCAFLSVK-----QFEELTGLBFLKLDRAKNKIAKTNNKKKEFE 126  
 DB 227 WEEHQKFSASCDIFKLGKLBKWTNEALMGLARITIMQYKGVSWLIDELKENRIDE 286  
 QY 127 TAK 129  
 DB 287 IIK 289

RESULT 4  
 A55478  
 neuronal apoptosis inhibitory protein - human  
 N:Alternate names: NAIP  
 C:Species: Homo sapiens (man)  
 C:Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 02-Feb-2001  
 C:Accession: A55478  
 R:ROY, N.; Mahadevan, M.S.; McLean, M.; Shuttler, G.; Varaghi, Z.; Parahani, R.; Baird, S.  
 Cell 80, 167-178, 1995  
 A:Title: The gene for neuronal apoptosis inhibitory protein is partially deleted in indiv  
 A:Reference number: A55478; MUID:95112344; PMID:7813013  
 A:Accession: A55478  
 A:Molecule type: mRNA  
 A:Residues: 1-1232 <ROY>  
 A:Cross-references: GB:U19251  
 C:Genetics:  
 A:Gene: GDB:SMA@; SMA  
 A:Cross-references: GDB:120378; OMIM:600354; OMIM:253300  
 A:Map position: 5q12.2-5q13  
 C:Keywords: apoptosis; ATP; glycoprotein; nucleotide binding; P-loop; transmembrane prote  
 F:94-110/Domain: transmembrane #status predicted <TM1>  
 F:470-477/Region: nucleotide-binding motif A (P-loop)  
 F:479-496/Domain: transmembrane #status predicted <TM2>  
 F:476/Binding site: ATP (lys) #status predicted  
 F:618,632,823,923,1035/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.9%; Score 155; DB 2; Length 1232;  
 Best Local Similarity 40.5%; Pred. No. 1.9e-05;  
 Matches 34; Conservative 15; Mismatches 29; Indels 6; Gaps 3;  
 QY 15 KDHRISTFKNWPF--LEGACCTPERMAEAGFIHCPTENEPDLAOCFFCFKELEGWEPDDDDP 73  
 DB 159 BEARLASFRNPFYVQG--ISPCVLSEAGFVF--TGKQDTVQCFSGCGCLGNWEEGDDP 213  
 QY 74 IEHKKHSSGCAFLSVKQFELT 97  
 DB 214 WKEHAKWPFKCFELRSKSSSEIT 237

RESULT 5  
 T03183  
 probable apoptosis inhibitor - Chilo iridescent virus  
 C:Species: Chilo iridescent virus  
 C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 02-Sep-2000  
 C:Accession: T03183  
 R:Bahr, U.; Tidona, C.A.; Darai, G.  
 Virus Genes 15, 235-245, 1997  
 A:Title: The DNA sequence of Chilo iridescent virus between the genome coordinates 0.101  
 A:Reference number: Z14834; MUID:98141693; PMID:9482589  
 A:Accession: T03183  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-208 <BAH>  
 A:Cross-references: EMBL:AF003534; NID:g2738385; PIDN:AAB94481.1; PID:g2738454  
 C:Superfamily: RING finger homology  
 F:159-202/Domain: RING finger homology <RNF>

Query Match 19.8%; Score 154; DB 2; Length 208;  
 Best Local Similarity 32.5%; Pred. No. 3.8e-06;  
 Matches 39; Conservative 23; Mismatches 38; Indels 20; Gaps 7;  
 QY 12 PFLK-DHRISTFKNWPFLEGCACCTPERMAEAGFIHCPTENEPDLAOCFFCFKELEGWEPD 70  
 DB 33 PPSADEKLNSQNNFPI--QLLPSEKESRAGFIYL---NIGDQVQCFYCDLKLKWKRS 87  
 QY 71 DDPIEHEKHSS-----GCAFLSVKQKQFELTGLBFLKLDREKRA-KNKIAKTNNKKKEFE 125  
 DB 88 DNPFEHKKHKTQDLKINCLFV-----KSIBFDNFVKHSESCFQNPFI---TNNINQOLD 138  
 RESULT 6

T37471  
apoptosis inhibitor homolog T27F2.3 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C/Accession: T37471; T25380  
R/Uren, A.G.  
submitted to the EMBL Data Library, January 1997  
A/Description: C. elegans IAP homologue.  
A/Reference number: Z21708  
A/Accession: T37471  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-155 <URE>  
A/Cross-references: EMBL:U85911; PIDN: AAB94330.1  
R/Lenhard, N.  
submitted to the EMBL Data Library, June 1996  
A/Reference number: Z20025  
A/Accession: T25380  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-155 <WIL>  
A/Cross-references: EMBL:Z74045; NID:e1062212; PIDN:CAA98553.1; GSPDB:GN00023; CESP:T27F2  
A/Experimental source: clone T27F2  
C/Genetics:  
A/Gene: T27F2.3  
A/Map position: 5  
A/Introns: 41/3

Query Match 18.7%; Score 145.5; DB 2; Length 155;  
Best Local Similarity 37.1%; Pred. No. 1.5e-05;  
Matches 36; Conservative 19; Mismatches 33; Indels 9; Gaps 6;  
QY 13 FLKDRIHSTFKNWFLE--GCATPERMAEAGFIHCPTNEPDLAQCFPCFKELEGWEPDDPI 70  
Db 16 FYKD-RLMTFKNFEDRDPDAKTSQVAQAF-YC---TGFSQKCAFCNKELD-FDPE 69  
QY 71 DPIEBKHKSSGCAFLSV-KQCFEELTLGEFLKIDR 106  
Db 70 DDPWEHTKRDPECFEVRICKLDDSELINDVRLSQ 106

RESULT 7  
T42628  
neuronal apoptosis inhibitory protein 2 - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 11-Jan-2000  
R/Yaraghi, Z.; Diez, E.; Gros, P.; Mackenzie, A.  
Mamm. Genome 10, 761-763, 1999  
A/Title: CDNA cloning and the 5' genomic organization of Naip2, a candidate gene for murine  
A/Reference number: Z22179; MUID:99315342; PMID:10384056  
A/Accession: T42628  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1447 <YAS>  
A/Cross-references: EMBL:AF102871; NID:g3860228; PID:g3860229; PIDN:AAC73002.1  
C/Genetics:  
A/Gene: Naip2

Query Match 18.3%; Score 142; DB 2; Length 1447;  
Best Local Similarity 37.3%; Pred. No. 0.00029;  
Matches 31; Conservative 13; Mismatches 35; Indels 4; Gaps 2;

QY 15 KDHRISTFKNWFLEGCATPERMAEAGFIHCPTNEPDLAQCFPCFKELEGWEPDDPI 74  
Db 159 EEARLESFEDWPFY-AHGTSPRVLSAAGFV---TKRDTVCFCSCGGCLGNWEEGDDPW 214  
QY 75 EEKHKSSGCAFLSVKQFEELT 97  
Db 215 KGHAKWFPKCEFLQSKSPEEIT 237

RESULT 8

A45679  
inhibitor-of-apoptosis polypeptide (IAP) - Cydia pomonella granulosis virus CpGV  
C/Species: Cydia pomonella granulosis virus CpGV  
C/Date: 21-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 15-Sep-2000  
C/Accession: A45679  
R/Crook, N.E.; Clem, R.J.; Miller, L.K.  
J. Virol. 67, 2168-2174, 1993  
A/Title: An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif.  
A/Reference number: A45679; MUID:93188168; PMID:8445726  
A/Accession: A45679  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-275 <CRO>  
A/Cross-references: GB:L05494; NID:g289583; PIDN:AAA43835.1; PID:g289584  
A/Note: sequence extracted from NCBI backbone (NCBIN:127014, NCBIPI:127015)  
C/Superfamily: viral apoptosis inhibitor IAP; RING finger homology

Query Match 17.5%; Score 136.5; DB 2; Length 275;  
Best Local Similarity 36.4%; Pred. No. 0.00016;  
Matches 28; Conservative 14; Mismatches 24; Indels 11; Gaps 3;

QY 18 RISTFKNWFLEGGAC---TPERMAEAGFIHCPTNEPDLAQCFPCFKELEGWEPDDPI 74  
Db 111 RVKSFHNWP-----RCMKQRPQVADAGFFY---TGYGDNTKFCYCDGGLKDWEPEDVPV 162  
QY 75 EEKHKSSGCAFLSVKX 91  
Db 163 EQVRWFDRCAVYVLVK 179

RESULT 9  
JC7568  
kidney inhibitor of apoptosis protein - human  
C/Species: Homo sapiens (man)  
C/Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
R/Lin, J.H.; Deng, G.; Huang, Q.; Morser, J.  
Biochem. Biophys. Res. Commun. 279, 820-831, 2000  
A/Title: KIAIP, a novel member of the inhibitor of apoptosis protein family.  
A/Reference number: JC7568; MUID: 21092523; PMID:11162435  
A/Contents: Fetal kidney  
A/Accession: JC7568  
A/Molecule type: mRNA  
A/Residues: 1-298 <LIN>  
C/Comment: This protein, a new member of the inhibitor of apoptosis protein family, plays

C/Genetics:  
A/Gene: kiap  
A/Map position: 20q13.3  
C/Keywords: apoptosis  
Query Match 17.5%; Score 136.5; DB 2; Length 298;  
Best Local Similarity 38.6%; Pred. No. 0.00018;  
Matches 27; Conservative 9; Mismatches 29; Indels 5; Gaps 2;

QY 18 RISTFKNWFLEGCATPERMAEAGFIHCPTNEPDLAQCFPCFKELEGWEPDDPI 77  
Db 90 RLASFYDWPLT--AEVPELLAAGFFH---TGQDKVRQCFYCGGLQSWKRGDDPWTEH 144  
QY 78 KKHSSGCAFL 87  
Db 145 AKWFPSQFL 154

RESULT 10  
TI0304

inhibitor of apoptosis protein 3 - Orgyia pseudotsugata nuclear polyhedrosis virus  
C/Species: Orgyia pseudotsugata nuclear polyhedrosis virus, OpMNPV  
C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 15-Sep-2000  
C/Accession: TI0304  
R/Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohrmann, G.F.  
Virology 229, 381-399, 1997  
A/Title: The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear polyhedrosis

```

F;567-611/Domain: RING finger homology <RNG>

Query Match          16.8%; Score 131; DB 2; Length 618;
Best Local Similarity 34.1%; Pred. No. 0.0011;
Matches 29; Conservative 12; Mismatches 34; Indels 10; Gaps 3;

QY 5 TLPPAWQFLKDHRISTFKNWP--FLEGCACTPERMAEAGFHCHPTENEPDLAQCFFCFK 62
DB 174 TUPYVAMSTEARLTHVMPLTFL-----SPSELAERAGFYI---GPGDRVACFACGG 225

QY 63 ELEGWEPDDDDPIEHHKHSSGCAFL 87
DB 226 KLSNWKPKDAMSEHRRHFPNCPFL 250

RESULT 13
T30489
apoptosis inhibitor - Lymantria dispar nuclear polyhedrosis virus
C;Species: Lymantria dispar nuclear polyhedrosis virus, LQMPV
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T30489
R;Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohrer,
Virology 253, 17-34, 1999
A;Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria di
A;Reference number: Z20836; MUID:99124785; PMID:9887315
A;Accession: T30489
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-155 <KUZ>
A;Cross-references: EMBL:AF081810; PIDN:AAC70325.1

Query Match          16.7%; Score 130; DB 2; Length 155;
Best Local Similarity 25.0%; Pred. No. 0.00033;
Matches 31; Conservative 29; Mismatches 44; Indels 20; Gaps 5;

QY 18 RISTKWNPFLEGCACTPERMAEAGFHCHPTENEPDLAQCFFCFKELGWEFPDDPIEHH 77
DB 7 RLASFRNNSAVD--APAPAEALAHAGF-YC--ANRQDFVKCAICHIEIGNWIGSDAMSDH 61

QY 78 KXHSSGCAFLSVKQFEEILTGEFLKDRERAKNKIAKETNNKKKEFEETAKKVRRAIEQ 137
DB 62 KRYSPACRFVC-----ELIK-----RPVSPVERDDDDDEEDSAAEPARGSEL 106

QY 138 LAAM 141
DB 107 LCSV 110

RESULT 14
JC5964
apoptosis inhibitor - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 02-Sep-2000
C;Accession: JC5964
R;Stehlik, C.; de Martin, R.; Binder, B.R.; Lipp, J.
Biochem. Biophys. Res. Commun. 243, 827-832, 1998
A;Title: Cytokine induced expression of porcine inhibitor of apoptosis protein (iap) fami
A;Reference number: JC5964; MUID:98162622; PMID:9501011
A;Accession: JC5964
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-358 <STE>
A;Cross-references: GB:U9142; NID:g2957174; PIDN:AAC39171.1; PID:g2957175
C;Superfamily: RING finger homology
F;307-351/Domain: RING finger homology <RRN>

Query Match          16.5%; Score 128.5; DB 2; Length 358;
Best Local Similarity 38.8%; Pred. No. 0.001;
Matches 31; Conservative 10; Mismatches 28; Indels 11; Gaps 4;

QY 15 KDHRISTFKNWP--FLEGCACTPERMAEAGFHCHPTENEPDLAQCFFCFKELGWEPPDD 72
DB 5 KD-RLITFQWMLTFL-----SPADLAKAGFYI---GPGDRVACFACGGKLSNWEPKDD 55

```



This Page Blank (uspto)

THIS PAGE BLANK (uspto)  
PENDING

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 14:10:30 ; Search time 14 Seconds  
(without alignments)  
528.140 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 778

Sequence: 1 MGAPTLPPAWQPFKDRHS.....EFETAKVRRRAIEQLAAMD 142

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	100.0	142	1 BIR5_HUMAN	O15392 homo sapien
2	660	84.8	140	1 BIR5_MOUSE	O70201 mus musculus
3	660	84.8	142	1 BIR5_RAT	O91jy7 rattus norv
4	163	21.0	997	1 BIR1_SCHPO	O14064 schizosacch
5	160	20.6	4829	1 BIR6_MOUSE	O9nr09 homo sapien
6	155	19.9	1403	1 BIR2_HUMAN	Q13075 homo sapien
7	154	19.8	208	1 ZFP_IRV6	P47732 chilo iride
8	145	18.6	1403	1 BIR6_MOUSE	O9jib6 mus musculus
9	143	18.4	1402	1 BIR6_MOUSE	O9jib3 mus musculus
10	143	18.4	1403	1 BIR6_MOUSE	O9qwk5 mus musculus
11	143	18.4	1403	1 BIR6_MOUSE	O9r016 mus musculus
12	142	18.3	1447	1 BIR6_MOUSE	O9qk4 mus musculus
13	136.5	17.5	275	1 IAP_GVCP	P41436 cydia pomon
14	136.5	17.5	298	1 BIR7_HUMAN	O96ca5 homo sapien
15	133.5	17.2	268	1 IAP3_NVOP	P41437 orgyia pseu
16	133.5	17.2	496	1 BIR4_MOUSE	O60989 mus musculus
17	133	17.1	496	1 BIR4_RAT	O9r016 rattus norv
18	131	16.8	618	1 BIR2_HUMAN	O13490 homo sapien
19	130	16.7	236	1 BIR8_HUMAN	O96009 homo sapien
20	129.5	16.6	612	1 BIR2_MOUSE	O62210 mus musculus
21	129	16.6	236	1 BIR8_PANTR	O95702 pan troglod
22	128.5	16.5	358	1 PIAP_PIG	O62640 sus scrofa
23	125.5	16.1	497	1 BIR4_HUMAN	P98170 homo sapien
24	125	16.1	236	1 BIR8_GORGO	O95m71 gorilla gor
25	125	16.1	600	1 BIR3_MOUSE	O08863 mus musculus
26	122	15.7	604	1 BIR3_HUMAN	Q13489 homo sapien
27	119.5	15.4	438	1 IAP1_DROME	Q24306 drosophila
28	118.5	15.2	611	1 BIR_CHECK	O90660 gallus gall
29	118	15.2	498	1 IAP2_DROME	Q24307 drosophila
30	106.5	13.7	954	1 BIR1_YEAST	P47134 saccharomyc
31	99.5	12.8	275	1 IAP1_NVOP	O10296 orgyia pseu
32	83	10.7	434	1 ADA2_YEAST	O02336 saccharomyc
33	82	10.5	135	1 LECG_BITAR	O92960 bitis ariet

#### ALIGNMENTS

##### RESULT 1

##### BIR5\_HUMAN

ID	BIR5_HUMAN	STANDARD;	PRT;	142 AA.
AC	O15392; Q9P2W8;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Baculoviral IAP repeat-containing protein 5 (Apoptosis inhibitor survivin) (Apoptosis inhibitor 4).			
GN	BIRC5 OR API4 OR IAP4.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM ALPHA).			
RX	MEDLINE=97399388; PubMed=9256286;			
RA	Ambrosini G., Adida C., Altieri D.C.;			
RT	"A novel anti-apoptosis gene, survivin, expressed in cancer and lymphoma."			
RL	Nat. Med. 3:917-921 (1997).			
RL	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM ALPHA).			
RC	TISSUE=Pancratic carcinoma;			
RA	Uren A.G., Vaux D.L.;			
RT	"Mammalian inhibitor of apoptosis (IAP) homolog."			
RL	Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM BETA).			
RA	Kageyama H., Islam A., Takayasu H., Nakagawara A.;			
RT	"An isoform of survivin (survivin-beta) which has 23 amino acids insertion into the BIR domain."			
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A. (ISOFORM ALPHA).			
RC	TISSUE=Lung, and Muscle;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K., Altschul S.F., Zeeberg B., Buetow K.H., Wang J., Hsieh F., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marzina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettner M., Maman A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."			

34	82	10.5	135	1 LECG_CROAT	P21963 crotalus at
35	81.5	10.5	286	1 IAP1_NVAC	P41435 autographa
36	80	10.3	213	1 R29_MOUSE	P97762 mus musculus
37	80	10.3	1391	1 M872_DROHY	Q08696 drosophila
38	79	10.2	221	1 R29_HUMAN	Q28486 homo sapien
39	79	10.2	226	1 YB34_METUA	Q58534 methanococc
40	78	10.0	1052	1 BULB_MOUSE	Q92180 mus musculus
41	77.5	10.0	454	1 TIG_BORBU	O51555 borrelia bu
42	77.5	10.0	489	1 DM2_MOUSE	P23804 mus musculus
43	76.5	9.8	361	1 CB45_MOUSE	Q61112 mus musculus
44	76	9.8	1818	1 HMW2_MYCPN	P75471 mycoplasma
45	75.5	9.7	180	1 GLUC_RAT	P06883 rattus norv

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [5]
RP FUNCTION.
RX MEDLINE=9075336; PubMed=9859933;
RA Li F., Ambrosini G., Chu E.Y., Plescia J., Tognin S., Marchisio P.C.,
RA Altieri D.C.;
RT "Control of apoptosis and mitotic spindle checkpoint by survivin.";
RL Nature 396:580-584 (1998).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.71 ANGSTROMS) OF ISOFORM ALPHA.
RX MEDLINE=20403315; PubMed=10949039;
RA Chantalat L., Skoufias D.A., Klemm J.P., Jung B., Dideberg O.,
RA Margolis R.L.;
RT "Crystal structure of human survivin reveals a bow tie-shaped dimer
RT with two unusual alpha-helical extensions.";
RL Mol. Cell 6:183-189 (2000).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.58 ANGSTROMS) OF ISOFORM ALPHA.
RX MEDLINE=20336902; PubMed=10876248;
RA Verdecia M.A., Huang H., Duttil E., Kaiser D.A., Hunter T., Noel J.P.;
RA Tognin S., Marchisio P.C., Altieri D.C.;
RT "Structure of the human anti-apoptotic protein survivin reveals a
RT dimeric arrangement.";
RL Nat. Struct. Biol. 7:602-608 (2000).
RN [8]
RP PHOSPHORYLATION OF THR-34.
RX MEDLINE=20542086; PubMed=11069302;
RA O'Connor D.S., Groesman D., Plescia J., Li F., Zhang H., Villa A.,
RA Tognin S., Marchisio P.C., Altieri D.C.;
RT "Regulation of apoptosis at cell division by p34cdc2 phosphorylation
RT of survivin.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:13103-13107 (2000).
CC -!- FUNCTION: May play a role in neoplasia. May counteract a default
CC induction of apoptosis in G2/M phase. Interacts with tubulin.
CC Inhibitor of caspase-3 and caspase-7.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Alpha;
CC IsoId=O15392-1; Sequence=Displayed;
CC Name=Beta;
CC IsoId=O15392-2; Sequence=VSP_002454;
CC -!- TISSUE SPECIFICITY: Expressed only in fetal kidney and liver, and
CC to lesser extent, lung and brain. Abundantly expressed in
CC adenocarcinoma (lung, pancreas, colon, breast, and prostate) and
CC in high-grade lymphomas.
CC -!- SIMILARITY: Contains 1 BIR repeat.
CC -!- SIMILARITY: Belongs to the IAP family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/ebis-sib.ch
CC or send an email to license@ebi-sib.ch).
CC
CC -----
CC EMBL; U75285; AAC51660.1; -
CC EMBL; AF077350; AAD34226.1; -
CC EMBL; AB028869; BAA93676.1; -
CC EMBL; BC008718; AAH08718.1; -
CC EMBL; BC034148; AAH34148.1; -
CC FDB; 1231; 03-JAN-01.
CC FDB; 1F3H; 16-MAY-01.
CC Genew; HGNC:593; BIRC5.
CC MIM; 603352; -
CC GO; GO:0005876; C:spindle microtubule; TAS.
CC GO; GO:0008189; P:apoptosis inhibitor activity; TAS.
CC GO; GO:0006916; P:anti-apoptosis; TAS.
CC GO; GO:0000086; P:G2/M transition of mitotic cell cycle; TAS.
CC InterPro; IPR001370; BIR.
CC Pfam; PF00653; BIR; 1.
CC SMART; SM00238; BIR; 1.

```

```

DR PROSITE; PS01282; BIR_REPEAT_1; FALSE_NEG.
DR PROSITE; PS0143; BIR_REPEAT_2; 1.
KW Apoptosis; Thiol protease inhibitor; Alternative splicing;
KW Metal-binding; Zinc; Phosphorylation; 3D-structure.
FT REPEAT 15 87
FT METAL 57 57
FT METAL 60 60
FT METAL 77 77
FT METAL 84 84
FT METAL 84 84
FT MOD_RES 34 34
FT VARSPLIC 74 74
FT I -> IGPWTAYACNTSLGGGRITR (in isoform
FT Beta).
FT FTIC=VSP_002454.
FT C->A: LOSS OF CYTOPROTECTION.
FT MUTAGEN 84 84
FT HELIX 8 13
FT HELIX 15 20
FT TURN 21 21
FT TURN 29 30
FT HELIX 35 40
FT TURN 41 42
FT STRAND 43 45
FT STRAND 55 57
FT TURN 58 60
FT STRAND 63 64
FT TURN 69 70
FT HELIX 73 80
FT TURN 82 83
FT HELIX 85 87
FT TURN 88 88
FT HELIX 93 95
FT STRAND 97 97
FT HELIX 98 139
SQ SEQUENCE 142 AA; 16389 MW; 9E7CADCDF2822286 CRC64;
Query Match 100.0%; Score 778; DB 1; Length 142;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGAPTLPPAQWFLKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTEHPDLAQCFPC 60
Db 1 MGAPTLPPAQWFLKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTEHPDLAQCFPC 60
QY 61 FKELEGWEDDDPIEHHKGGSGCAFLSVKKQFELTIGEFKLDREKAKKIATETNNK 120
Db 61 FKELEGWEDDDPIEHHKGGSGCAFLSVKKQFELTIGEFKLDREKAKKIATETNNK 120
QY 121 KKEFEETAKKVRRAIEQLAAMD 142
Db 121 KKEFEETAKKVRRAIEQLAAMD 142
RESULT 2
BIR5_MOUSE STANDARD; PRT; 140 AA.
AC O70201; Q923F7; Q9WU53; Q9WU54;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Baculoviral IAP repeat-containing protein 5 (Apoptosis inhibitor
DE survivin) (Apoptosis inhibitor 4) (TIAP).
GN BIRC5 OR API4 OR IAP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Uren A.G., Vaux D.L.;
RT "Mammalian inhibitor of apoptosis (IAP) homolog.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Embryo;

```



RA Kobayashi K., Otaki M., Ogasawara T., Tokuhisa T.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
 RX MEDLINE=20129603; PubMed=10666222;  
 RA Conway E.M., Pollefeyt S., Cornelissen J., DeBaere I.,  
 RA Steiner-Mosonyi M., Ong K., Baens M., Colleen D., Schuh A.C.;  
 RT "three differentially expressed survivin cDNA variants encode  
 RL proteins with distinct antiapoptotic functions.";  
 RL Biol 95:1435-1442(2000).  
 [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan M., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McWan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RL human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 7-118, AND PHOSPHORYLATION  
 OF THR-34.  
 RX MEDLINE=20403314; PubMed=10949038;  
 RA Muchmore S.W., Chen J., Jakob C., Zakula D., Matayoshi E.D., Wu W.,  
 RA Zhang H., Li F., Ng S.C., Altieri D.C.;  
 RT "Crystal structure and mutagenic analysis of the  
 RL inhibitor of cell apoptosis protein survivin.";  
 RL Mol. Cell 6:173-182(2000).  
 CC -1- FUNCTION: May play a role in neoplasia. May counteract a default  
 CC induction of apoptosis in G2/M phase. Interacts with tubulin.  
 CC Inhibitor of caspase-3 and caspase-7 (By similarity).  
 CC -1- SUBUNIT: Homodimer; zinc-dependent.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1; Synonyms=Survivin 140;  
 CC IsoId=O70201-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Survivin 121;  
 CC IsoId=O70201-2; Sequence=VSP\_002457;  
 CC Name=3; Synonyms=Survivin 40;  
 CC IsoId=O70201-3; Sequence=VSP\_002455, VSP\_002456;  
 CC -1- SIMILARITY: Contains 1 BIR repeat.  
 CC -1- SIMILARITY: Belongs to the IAP family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF077349; AA034225.1; -  
 CC EMBL; AB013819; BA228266.1; -  
 CC EMBL; AF115517; AA026199.1; -  
 CC EMBL; AF115517; AA026200.1; -  
 CC EMBL; AF115517; AA026201.1; -  
 CC EMBL; BC004702; AA04702.1; -  
 CC PDB; 1M4M; 25-SEP-02.

DR MGI:1203517; Birc5.  
 DR GO; GO:0005737; C:cytoplasm; IDA.  
 DR GO; GO:0008189; F:apoptosis inhibitor activity; IMP.  
 DR GO; GO:0006916; P:anti-apoptosis; IMP.  
 DR InterPro; IPR001370; BIR.  
 DR Pfam; PF00653; BIR; 1.  
 DR SMART; SM00238; BIR; 1.  
 DR PROSITE; PS01282; BIR\_REPEAT\_1; FALSE\_NEG.  
 DR PROSITE; PS0143; BIR\_REPEAT\_2; 1.  
 KW Apoptosis; Thiol protease inhibitor; Alternative splicing;  
 KW Phosphorylation; Metal-binding; Zinc; 3D-structure.  
 FT REPEAT 15 87  
 FT MOD\_RES 34 34 PHOSPHORYLATION (BY CDC2).  
 FT METAL 57 57 ZINC 1.  
 FT METAL 60 60 ZINC 1.  
 FT METAL 76 76 ZINC 2.  
 FT METAL 77 77 ZINC 1.  
 FT METAL 80 80 ZINC 2.  
 FT METAL 84 84 ZINC 1.  
 FT VARSPLIC 38 40 MAE -> RGA (in isoform 3).  
 FT VARSPLIC 41 140 /FTid=VSP\_002455.  
 FT VARSPLIC 114 140 Missing (in isoform 3).  
 FT VARSPLIC 114 140 /FTid=VSP\_002456.  
 FT AKETNNKQKEFEETAKTTQSIQLAA -> VCMENKD  
 FT (in isoform 2).  
 FT /FTid=VSP\_002457.  
 SQ SEQUENCE 140 AA; 16297 MW; 26F5ABF501A5D83C CRC64;  
 Query Match 84.8%; Score 660; DB 1; Length 140;  
 Best Local Similarity 84.3%; Pred. No. 8.3e-51;  
 Matches 118; Conservative 12; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 MGAPLPPAWQFLLKDHRISTFKWPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFFC 60  
 DB 1 MGAPALPQWLYLYKNYRIATFKWPFLEDCACTPERMAEAGFIHCPTENEPDLAQCFFC 60  
 QY 61 FKELEGWPDPTIEHKKHSGGCAFLSVKKQFBEITLGEFLKLDRAKAKIATNNK 120  
 DB 61 FKELEGWPDPTIEHKKHSGGCAFLSVKKQFBEITLGEFLKLDRAKAKIATNNK 120  
 QY 121 KKEFEETAKVRRATEQLAA 140  
 DB 121 KKEFEETAKTTQSIQLAA 140  
 RESULT 3  
 BIR5\_RAT  
 ID BIR5\_RAT STANDARD; PRT; 142 AA.  
 AC Q9JH77;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Baculoviral IAP repeat-containing protein 5 (Apoptosis inhibitor  
 DE survivin).  
 GN BIRC5.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RA Chen D., Cao G., Chen J.;  
 RT "Molecular cloning and characterization of rat survivin.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: May play a role in neoplasia. May counteract a default  
 CC induction of apoptosis in G2/M phase. Interacts with tubulin.  
 CC Inhibitor of caspase-3 and caspase-7 (By similarity).  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Contains 1 BIR repeat.  
 CC -1- SIMILARITY: Belongs to the IAP family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; AF276775; AAF82586.1; --  
 CC HSSP; Q15392; 1E31.  
 CC InterPro; IPR001370; BIR.  
 CC Pfam; PF00653; BIR; 1.  
 CC SMART; SM00238; BIR; 1.  
 CC PROSITE; PS01282; BIR\_REPEAT\_1; FALSE\_NEG.  
 CC PROSITE; PS01443; BIR\_REPEAT\_2; 1.  
 CC Apoptosis; Thiol protease inhibitor; Metal-binding; Zinc;  
 KW Phosphorylation.  
 FT REPEAT 15 87 BIR.  
 FT METAL 57 57 ZINC (BY SIMILARITY).  
 FT METAL 60 60 ZINC (BY SIMILARITY).  
 FT METAL 77 77 ZINC (BY SIMILARITY).  
 FT METAL 84 84 ZINC (BY SIMILARITY).  
 FT MOD\_RES 34 34 PHOSPHORYLATION (BY CDC2)  
 FT (BY SIMILARITY).  
 SQ SEQUENCE 142 AA; 16692 MW; 99BCFAE15B4D0CAC CRC64;

Query Match 84.8%; Score 660; DB 1; Length 142;  
 Best Local Similarity 83.0%; Pred. No. 8.4e-51;  
 Matches 117; Conservative 13; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 MGAPILPAWQFLKXDRHISTFKWPFLEGGCACTPERMAEAGFHCPTEPNPDLAQCFFC 60  
 Db |||||  
 QY 1 MGATLPPIWQWYLDXDRHLYFKWPFLEDCSCPERMAEAGFHCPTEPNPDLAQCFFC 60  
 Db |||||  
 QY 61 FKELGWSPDDPIBEHKHSGGCAFLVKVQFELTGEFLKLDREKAKYKAKETNNK 120  
 Db |||||  
 QY 61 FKELGWSPDDPIBEHKHSGGCAFLVKVQFELTGEFLKLDREKAKYKAKETNNK 120  
 Db |||||  
 QY 121 KKEFEETAKKVRATEQLAAM 141  
 Db |||||  
 QY 121 KKEFEETRTVRSIEQLAAL 141  
 Db |||||

RESULT 4  
 BIR1 SCHPO  
 ID BIR1 SCHPO STANDARD; PRT; 997 AA.  
 AC Q14064; Q9USG4;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DE BIR1 protein (Chromosome segregation protein cut17).  
 DE BIR1 OR CUT17 OR PBH1 OR SPCC962.02C OR SPCC31B10.10C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 GN Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetes;  
 OC Schizosaccharomycetes;  
 OC NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.  
 RX MEDLINE=21439264; PubMed=11554922;  
 RA Morishita J., Matsusaka T., Goshima G., Nakamura T., Tatebe H.,  
 RA Yanagida M.,  
 RT "Bir1/Cut17 moving from chromosome to spindle upon the loss of  
 RT cohesion is required for condensation, spindle elongation and  
 RT repair."  
 RL Genes Cells 6:743-763 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., William R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Gaffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shipakowski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880 (2002).  
 RN [3]  
 RN CHARACTERIZATION.  
 RP MEDLINE=99398681; PubMed=10468581;  
 RX Uren A.G., Beilharz T., O'Connell M.J., Bugg S.J., van Driel R.,  
 RA Vaux D.L., Litgow T.;  
 RT "Role for yeast inhibitor of apoptosis (IAP)-like proteins in cell  
 RT division."  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:10170-10175 (1999).  
 RN [4]  
 RN CHARACTERIZATION.  
 RP MEDLINE=21850422; PubMed=11861551;  
 RX Rajagopalan S., Balasubramanian M.K.;  
 RA "Schizosaccharomyces pombe Birp, a nuclear protein that localizes to  
 RT kinetochores and the spindle midzone, is essential for chromosome  
 RT condensation and spindle elongation during mitosis."  
 RL Genetics 160:445-456 (2002).  
 RN [5]  
 RN FUNCTION.  
 RP MEDLINE=20035862; PubMed=10571085;  
 RX Rajagopalan S., Balasubramanian M.K.;  
 RA "S. pombe Phip: an inhibitor of apoptosis domain containing protein  
 RT is essential for chromosome segregation."  
 RL FEBS Lett. 460:187-190 (1999).  
 CC -!- FUNCTION: Seems to act in the pleiotropic control of cell  
 CC division. Has a role in chromosome segregation by recruiting  
 CC condensin and ark1 kinases to appropriate sites as the cell  
 CC progresses through mitosis.  
 CC -!- SUBCELLULAR LOCATION: Nuclear. Interacts with the outer  
 CC centromeric regions of the chromosomes during interphase. After  
 CC chromatid separation moves to the middle of the spindle.  
 CC -!- SIMILARITY: Contains 2 BIR repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; AB031034; BAA83415.1; --  
 CC DR EMBL; AL031323; CAA20434.1; --  
 CC DR EMBL; AL121859; CAB58376.1; --  
 CC DR PIR; T43523; T43523.  
 CC HSSP; Q13490; 1QBH.  
 CC GeneDB; SPombe; SPCC962.02C; --  
 CC InterPro; IPR001370; BIR.  
 CC Pfam; PF00653; BIR; 2.  
 CC SMART; SM00238; BIR; 2.  
 CC PROSITE; PS01282; BIR\_REPEAT\_1; FALSE\_NEG.

```

use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC ENBL; AP265555; AAF75772.1; -.
CC ENBL; AR033115; BAA86603.2; ALT_FRAME.
CC HSP; Q13490; IQBH.
CC Genew; HGNC:13516; BIRC6.
CC MIM; 605638; -.
CC GO; GO:0004840; F-ubiquitin conjugating enzyme activity; TAS.
CC GO; GO:0006916; P-anti-apoptosis; TAS.
CC InterPro; IPR001370; BIR.
CC InterPro; IPR000608; UBQ_conjugat.
CC Pfam; PF00653; BIR; 1.
CC Pfam; PF001179; UQ_con; 1.
CC ProDom; PD000461; UBQ_conjugat; 1.
CC SMART; SM00238; BIR; 1.
CC SMART; SM00212; UBCC; 1.
CC PROSITE; PS01282; BIR_REPEAT_1; FALSE_NEG.
CC PROSITE; PS0143; BIR_REPEAT_2; 1.
CC PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; FALSE_NEG.
CC PROSITE; PS01127; UBIQUITIN_CONJUGAT_2; 1.
CC KX Apoptosis; Thiol protease inhibitor; Ub1 conjugation pathway; Ligase.
FT REPEAT 256 330 BIR.
FT DOMAIN 4548 4676 UBIQUITIN-CONJUGATING.
FT BINDING 4597 4597 UBIQUITIN (BY SIMILARITY).
FT DOMAIN 1632 1640 POLY-ALA.
FT DOMAIN 2291 2291 L -> F (IN REF. 2).
FT CONFLICT 2646 2646 S -> T (IN REF. 2).
SQ SEQUENCE 4829 AA; 527604 MW; C67136A672CD3653 CRC64;

Query Match 20.6%; Score 160; DB 1; Length 4829;
Best Local Similarity 42.3%; Pred. No. 1.2e-05;
Matches 30; Conservative 10; Mismatches 31; Indels 0; Gaps 0;

Qy 17 HRISTKPNFFLEGCACTPERMAEAGFIHCPTENEPLDAQCFKFELEGWEPDDPIEE 76
Db 260 NRRETTTSFHVGYRWAQPDMAQAGFYHQPASSGDDRAMCFTCSVLCVWCEPTDEWSE 319
Qy 77 HKXHSAGCAPL 87
Db 320 HEHSPNCPFV 330

RESULT 6
BIR1_HUMAN
ID BIR1_HUMAN STANDARD; PRT; 1403 AA.
AC Q13075; O75857; Q13730; Q99796;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Baculoviral IAP repeat-containing protein 1 (Neuronal apoptosis
DE inhibitory protein).
GN BIRC1 OR NAIP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=95112344; PubMed=7813013;
RA Roy N., Mahadevan M.S., McLean M., Shuttler G., Yaraghi Z.,
RA Farahini R., Baird S., Besner-Johnston A., Lefebvre C., Kang X.,
RA Salih M., Aubry H., Tamai K., Guan X., Ioannou P., Crawford T.O.,
RA da Jong P.J., Surh L., Ikeda J., Korneluk R.G., Mackenzie A.;
RT "The gene for neuronal apoptosis inhibitory protein is partially
RT deleted in individuals with spinal muscular atrophy.";
RL Cell 80:167-178(1995).
RN [2]
RP SEQUENCE FROM N.A. AND REVISIONS.

```

PC TISSUE=Brain;  
RX MEDLINE=98163755; PubMed=9503025;  
RA Chen Q., Baird S.D., Mahadevan M., Besner-Johnston A., Farahani R.,  
RA Xuan J.-Y., Kang S.X., Lefebvre C., Ikeda J.-E., Korneluk R.G.,  
RA Mackenzie A.E.;  
RA "Sequence of a 131-kb region of 5q13.1 containing the spinal muscular  
RT atrophy candidate genes SMN and NAIP";  
PL Genomics 48:121-127(1998).  
RN [3].  
RP SEQUENCE OF 386-623 FROM N.A.  
RA der Steege G., Draaijers T.G., Grootsholten P.M., Osinga J.,  
RA Anzevino R., Velsona I., Brahe C., Scheffer H., van Ommen G.J.B.,  
RA Buys C.H.C.M.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [4].  
RP SEQUENCE OF 222-1403 FROM N.A.  
RA Jones K., Graves T., McPherson J.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
RN [5].  
RP FUNCTION.  
RC TISSUE=Liver;  
RX MEDLINE=96149249; PubMed=8552191;  
RA Liston P., Roy N., Tamai K., Lefebvre C., Baird S., Cherton-Horvat G.,  
RA Farahani R., McLean M., Ikeda J., Mackenzie A., Korneluk R.G.;  
RT "Suppression of apoptosis in mammalian cells by NAIP and a related  
RT family of IAP genes";  
RL Nature 379:349-353(1996).  
CC -!- FUNCTION: Prevents motor-neuron apoptosis induced by a variety of  
CC signals.  
CC -!- TISSUE SPECIFICITY: Expressed in motor neurons, but not in sensory  
CC neurons. Found in liver and placenta, and to a lesser extent in  
CC spinal cord.  
CC -!- DISEASE: Mutated or deleted forms of NAIP have been found in  
CC individuals with spinal muscular atrophy type I (SMA1)  
CC [MIM:253300]. SMA are fatal autosomal recessive disorders  
CC subclassified as type I (Werdnig-Hoffmann disease), type II  
CC (intermediate form), and type III (Wohlfart-Kugelberg-Welander  
CC disease) based upon the age of onset and clinical severity. These  
CC neurodegenerative disorders are characterized by degeneration of  
CC lower motor neurons, leading to progressive paralysis muscular  
CC atrophy. Concerns 1 in 6000 newborns.  
CC -!- SIMILARITY: Contains 3 BIR repeats.  
CC -!- SIMILARITY: Contains 1 NACHT domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U19251; AAC52045.1; -.  
CC EMBL; U80017; AAC52047.1; -.  
CC EMBL; U21913; AAA64504.1; -.  
CC EMBL; AC005031; AAC62261.1; -.  
CC HSSP; Q13490.1QBH.  
CC Genew; HGNC:7634; BIRC1.  
CC MIM; 600355; -.  
CC MIM; 253300; -.  
CC DR GO; 0008189; F:apoptosis inhibitor activity; TAS.  
CC DR GO; 0006916; P:anti-apoptosis; TAS.  
CC DR GO; 0007399; P:neurogenesis; TAS.  
CC DR InterPro; IPR003593; AAA\_ATPase.  
CC DR InterPro; IPR001370; BIR.  
CC DR InterPro; IPR007111; NACHT\_NTPase.  
CC DR Pfam; PF00653; BIR; 3.  
CC DR Pfam; PF05729; NACHT; 1.  
CC DR SMART; SM00382; AAA; 1.  
CC DR SMART; SM00238; BIR; 3.  
CC DR PROSITE; PS01282; BIR\_REPEAT\_1; 3.  
CC DR PROSITE; PS0143; BIR\_REPEAT\_2; 3.  
CC DR PROSITE; PS50837; NACHT; 1.

KW Apoptosis; Repeat. 127 BIR 1.  
FT REPEAT 60 BIR 2.  
FT REPEAT 159 BIR 3.  
FT REPEAT 278 BIR 3.  
FT DOMAIN 464 NACHT.  
FT CONFLICT 222 PK -> YR (IN REF. 4).  
FT CONFLICT 386 VP -> ST (IN REF. 3).  
FT CONFLICT 535 M -> V (IN REF. 3).  
FT CONFLICT 553 Y -> H (IN REF. 3).  
FT CONFLICT 1228 MISSING (IN REF. 4).  
SQ SEQUENCE 1403 AA; 159613 MW; 566304C154DA5E64 CRC64;  
  
Query Match 19.9%; Score 155; DB 1; Length 1403;  
Best Local Similarity 40.5%; Pred. No. 8.8e-06;  
Matches 34; Conservative 15; Mismatches 29; Indels 6; Gaps 3;  
  
QY 15 KDHRISTKFWPF-LEGACATPERMABAGFTHCTENEPDLAQCFKCELEGWEPDDDP 73  
Db 159 EARLASFRNWFVVGQ--ISPCVLSAGVF---TGKQDTVQCFCGCGCLGNWEEGDDP 213  
QY 74 IEHKKHSSGCAFLSVKQFBEIT 97  
Db 214 WKEHAKWPKCEFLRSKKSSEIT 237  
  
RESULT 7  
ZFP\_IRV6 STANDARD; PRT; 208 AA.  
AC P47732; O55770;  
DT 01-FEB-1996 (Rel. 33; Created)  
DT 10-OCT-2003 (Rel. 42; Last sequence update)  
DT 10-OCT-2003 (Rel. 42; Last annotation update)  
DE Zinc finger protein.  
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).  
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.  
OX NCBI\_TaxID:10488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92196996; PubMed=1549908;  
RA Handermann M., Schmitz P., Roosen-Wolff A.P., Raab K.,  
RA Sonntag K.C., Darai G.;  
RT "Identification and mapping of origins of DNA replication within the  
RT DNA sequences of the genome of insect iridescent virus type 6";  
RL Virus Genes 6:19-32(1992).  
RN [2]  
RP REVISIONS TO C-TERMINUS.  
RA Jakob N.J., Mueller K., Bahr U., Darai G.;  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Contains 1 BIR repeat.  
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF303741; AAB94481.1; -.  
CC HSSP; Q15392; IE31.  
CC InterPro; IPR001370; BIR.  
CC InterPro; IPR001841; Znf\_ring.  
CC Pfam; PF00653; BIR; 1.  
CC Pfam; PF00097; zf-C3HC4; 1.  
CC SMART; SM00238; BIR; 1.  
CC SMART; SM00184; RING; 1.  
CC PROSITE; PS01282; BIR\_REPEAT\_1; FALSE\_NEG.  
CC PROSITE; PS0143; BIR\_REPEAT\_2; 1.  
CC PROSITE; PS00518; ZF\_RING\_1; FALSE\_NEG.  
CC PROSITE; PS00089; ZF\_RING\_2; 1.  
KW Zinc-finger; Repeat.  
FT REPEAT 37 108 BIR.

```

FT DOMAIN          134      151      3 X 6 AA TANDEM REPEATS.
FT REPEAT         134      139      1.
FT REPEAT         140      145      2.
FT REPEAT         146      151      3.
FT ZN_FING        163      197      RING-TYPE.
SQ SEQUENCE       208 AA; 24142 MW; 41A2E2FC18833390 CRC64;

Query Match              19.8%; Score 154; DB 1; Length 208;
Best Local Similarity    32.5%; Pred. No. 1.4e-06;
Matches 39; Conservative 23; Mismatches 38; Indels 20; Gaps 7.

Qy      12 PFLK-DHRISFKNWPFLEGCACPTPERMAEAGFIHCPTENEPDLAQCFPCFKELGWEFPD 70
Db      33 PFASYDERLNSFNWPFI---QLLPKEQUSRAGFYIYLNIGDQQVCYSIDLKUEWKRS 87
Qy      71 DDIIEHHKKHGS-----GGAFSLVSKVKQFEELTGLGFLLKLDREKA-KNKIAKETNNKKKEFE 125
Db      88 DNPEEHKHGHKTQDLKINCLFV-----KSIEFDNFVKHSSSCFQNPI---TNINQDLD 138

RESULT 8
BIRF_MOUSE
ID BIRF_MOUSE STANDARD; PRT; 1403 AA.
AC QJUIBG; C09121; C09122; P81704;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Baculoviral IAP repeat-containing protein 1f (Neuronal apoptosis
inhibitory protein 6).
OS BIRC1F OR NAIP6 OR NAIP-RS4.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RM "The mouse region syntenic for human spinal muscular atrophy lies
within the Igln1 critical interval and contains multiple copies of Naip
exon 5."
RL Genomics 38:405-417(1996).
CC -1- FUNCTION: Prevents motor-neuron apoptosis induced by a variety of
signals.
CC -1- SIMILARITY: Contains 3 BIR repeats.
CC -1- SIMILARITY: Contains 1 NACHT domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR ENBL; AF242431; AAF82751.1; .
DR ENBL; U66327; AAC52975.1; .
DR HSSP; Q13490; IQBH.
DR MGD; MG1.I129822; Birg1f.
DR InterPro; IPRO01370; BIR.
DR InterPro; IPRO07111; NACHT_NTPase.
DR Pfam; PF00653; BIR_3.
DR Pfam; PF05729; NACHT; 1.
DR SMART; SM00238; BIR; 3.
DR PROSITE; PS01282; BIR_REPEAT 1; 2.

```

Tue Aug 17 05:54:58 2004

```

Query Match      18.4%; Score 143; DB 1; Length 1402;
Best Local Similarity 36.6%; Pred. No. 9.7e-05;
Matches 30; Conservative 13; Mismatches 35; Indels 4; Gaps 2;

Qy 15 KDRISTFKWPFLEGGACCTPERMAAGFTHCTENEPDLAQCFKCFKELEGWEPDDPI 74
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 159 EEARLESFEDWPFY-AGTSPRVLAAAGFV-----TGKRDIVQCFSCGSLNWEEDDPW 214
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 75 EEKHKSSGCAFLSVKKQFEEL 96
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 KEHAKWPKCFEFLQSKSSEI 236
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
BIRA MOUSE STANDARD; PRT; 1403 AA.
AC Q9DMK5; Q9JIB5; Q9R017;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Baculoviral IAP repeat-containing protein 1a (Neuronal apoptosis
DE inhibitory protein 1).
GN BIRC1A OR NAIP1 OR NAIP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98360097; PubMed=10958627;
RX Huang S., Scharf J.M., Gowney J.D., Endrizzi M.G., Dietrich W.F.;
RT "The mouse Naip gene cluster on Chromosome 13 encodes several distinct
RT functional transcripts.";
RL Mamm. Genome 10:1032-1035(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=20414747; PubMed=10958627;
RX Endrizzi M.G., Hadinoto V., Gowney J.D., Miller W., Dietrich W.F.;
RT "Genomic sequence analysis of the mouse Naip gene array.";
RL Genome Res. 10:1095-1102(2000).
CC -!- SIMILARITY: Contains 3 BIR repeats.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF007769; AAB69223.1; -
DR EMBL; AF135491; AAD56763.1; -
DR EMBL; AF242432; AAF82752.1; -
DR HSSP; Q13430; 1QBH.
DR MGD; MGI:1298223; Birc1a.
DR InterPro; IPR003593; AAA_NTPase.
DR InterPro; IPR001370; BIR.
DR InterPro; IPR007111; NACHT_NTPase.
DR Pfam; PF00653; BIR; 3.
DR Pfam; PF05729; NACHT; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00238; BIR; 3.
DR PROSITE; PS01282; BIR_REPEAT_1; 1.

```

```

DR PROSITE; PS50143; BIR_REPEAT_2; 3.
DR PROSITE; PS50837; NACHT; 1.
KW Apoptosis; Repeat; Multigene family.
FT REPEAT 60
FT REPEAT 127
FT REPEAT 159
FT REPEAT 227
FT REPEAT 278
FT REPEAT 345
FT DOMAIN 464
FT DOMAIN 758
FT CONFLICT 343
FT CONFLICT 359
FT CONFLICT 359
FT CONFLICT 624
FT CONFLICT 1092
FT CONFLICT 1092
FT CONFLICT 1116
FT CONFLICT 1123
FT CONFLICT 1129
FT CONFLICT 1140
FT CONFLICT 1269
FT CONFLICT 1403 AA; B31630259595EE67 CRC64;
SQ SEQUENCE 1403 AA; 158692 MW; B31630259595EE67 CRC64;

Query Match      18.4%; Score 143; DB 1; Length 1403;
Best Local Similarity 36.6%; Pred. No. 9.7e-05;
Matches 30; Conservative 13; Mismatches 35; Indels 4; Gaps 2;

Qy 15 KDRISTFKWPFLEGGACCTPERMAAGFTHCTENEPDLAQCFKCFKELEGWEPDDPI 74
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 159 EEARLESFEDWPFY-AGTSPRVLAAAGFV-----TGKRDIVQCFSCGSLNWEEDDPW 214
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 75 EEKHKSSGCAFLSVKKQFEEL 96
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 KEHAKWPKCFEFLQSKSSEI 236
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
BIRE MOUSE STANDARD; PRT; 1403 AA.
AC Q9R016; Q09121; Q09122; P81703; Q9R029;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Baculoviral IAP repeat-containing protein 1e (Neuronal apoptosis
DE inhibitory protein 5).
GN BIRC1E OR NAIP5 OR NAIP-RS3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99431676; PubMed=10501978;
RA Huang S., Scharf J.M., Gowney J.D., Endrizzi M.G., Dietrich W.F.;
RT "The mouse Naip gene cluster on Chromosome 13 encodes several distinct
RT functional transcripts.";
RL Mamm. Genome 10:1032-1035(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RC MEDLINE=99417674; PubMed=10486205;
RX Endrizzi M., Huang S., Scharf J.M., Kelter A.R., Wirth B.,
RX Kunkel L.M., Dietrich W.F.;
RA "The mouse region syntenic for human spinal muscular atrophy lies
RA within the Lgn1 critical interval and contains multiple copies of Naip
RA exon 5.";
RL Genomics 38:405-417(1996).
CC -!- FUNCTION: Prevents motor-neuron apoptosis induced by a variety of
CC signals.

```

```

CC -!- SIMILARITY: Contains 3 BIR repeats.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF135492; AAD56764.1; -
DR EMBL; AF131205; AAD56760.1; -
DR EMBL; U66326; AAC52974.1; -
DR HSSP; Q13490; IQBH.
DR MGD; MGI:1298220; Bircle.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 3.
DR Pfam; PF05729; NACHT; 1.
DR SMART; SM00238; BIR; 3.
DR PROSITE; PS01282; BIR_REPEAT_1; 2.
DR PROSITE; PS0143; BIR_REPEAT_2; 3.
DR PROSITE; PS0837; NACHT; 1.
KW Apoptosis; Repeat; Multigene family.
FT REPEAT 60 127
FT REPEAT 159 227
FT REPEAT 278 345
FT DOMAIN 464 759
FT CONFLICT 92 92 K -> R (IN REF. 1).
FT CONFLICT 144 144 S -> R (IN REF. 1).
FT CONFLICT 242 242 S -> G (IN REF. 2).
FT CONFLICT 472 472 T -> A (IN REF. 2).
FT CONFLICT 516 516 A -> D (IN REF. 2).
FT CONFLICT 521 521 A -> T (IN REF. 2).
FT CONFLICT 533 533 V -> A (IN REF. 2).
FT CONFLICT 538 538 S -> I (IN REF. 2).
FT CONFLICT 1092 1092 E -> L (IN REF. 2).
FT CONFLICT 1129 1129 H -> L (IN REF. 2).
FT CONFLICT 1137 1137 R -> Q (IN REF. 2).
FT CONFLICT 1242 1242 V -> I (IN REF. 2).
FT CONFLICT 1276 1276 D -> N (IN REF. 2).
SQ SEQUENCE 1403 AA; 159695 MW; B27F645043BCEC42 CRC64;

Query Match 18.4%; Score 143; DB 1; Length 1403;
Best Local Similarity 36.8%; Pred. No. 9.7e-05;
Matches 30; Conservative 13; Mismatches 35; Indels 4; Gaps 2;

QY 15 KDRISTFKWPFLEGACCTPERMAGFTHCPTENEPLAQCFPFKELEGWEPDDDFI 74
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 159 EEARLESFEDWPFV-AHGTSFRLVSAAGVFV---TGKRDVTQCFGCGSLGNWEGGDFW 214

QY 75 EEHKXSSGCAFLSVKQFEEL 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 KEHAKWPFKCEFLQSKKSBEI 236

RESULT 12
BIR_MOUSE STANDARD; PRT; 1447 AA.
AC Q9OUK4; Q9OUK4; Q9RO30;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Baculoviral IAP repeat-containing protein 1b (Neuronal apoptosis
DE inhibitory protein 2).
GN BIRC1B OR NAIP2 OR NAIP-RS6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=99431676; PubMed=10501978;
RA Huang S., Scharf J.M., Gowney J.D., Endrizzi M.G., Dietrich W.F.;
RT "The mouse Naip gene cluster on Chromosome 13 encodes several distinct
RT functional transcripts.";
RL Mamm. Genome 10:1032-1035(1999).
RL [2]
RP SEQUENCE FROM N.A.
RP MEDLINE=99315342; PubMed=10384056;
RA Yargali Z., Diez E., Gros P., Mackenzie A.;
RT "cDNA cloning and the 5' genomic organization of Naip2, a candidate
RT gene for murine Legionella resistance.";
RL Mamm. Genome 10:761-763(1999).
RL [3]
RP SEQUENCE FROM N.A.
RP STRAIN=129/SV;
RX MEDLINE=99417674; PubMed=10486205;
RA Endrizzi M., Huang S., Scharf J.M., Kelter A.R., With B.,
RA Kunkel L.M., Miller W., Dietrich W.F.;
RT "Comparative sequence analysis of the mouse and human Igln1/SMa
RT interval.";
RL Genomics 60:137-151(1999).
RL [4]
RP SEQUENCE OF 82-168 FROM N.A.
RX STRAIN=129/SV;
RX MEDLINE=97131520; PubMed=8975718;
RA Scharf J.M., Damron D., Frisella A., Bruno S., Beggs A.H.,
RA Kunkel L.M., Dietrich W.F.;
RT "The mouse region syntenic for human spinal muscular atrophy lies
RT within the Igln1 critical interval and contains multiple copies of Naip
RT exon 5.";
RL Genomics 38:405-417(1996).
CC -!- FUNCTION: Prevents motor-neuron apoptosis induced by a variety of
CC signals.
CC -!- SIMILARITY: Contains 3 BIR repeats.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF135489; AAD56761.1; -
DR EMBL; AF135490; AAD56762.1; -
DR EMBL; AF102871; AAC73002.1; -
DR EMBL; AF131205; AAD56759.1; -
DR EMBL; U66329; AAC52977.1; -
DR PIR; T42628; T42628.
DR HSSP; Q13490; IQBH.
DR MGD; MGI:1298226; Birc1b.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001370; BIR.
DR InterPro; IPR007111; NACHT_NTPase.
DR Pfam; PF00653; BIR; 3.
DR Pfam; PF05729; NACHT; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00238; BIR; 3.
DR PROSITE; PS01282; BIR_REPEAT_1; 2.
DR PROSITE; PS0143; BIR_REPEAT_2; 3.
DR PROSITE; PS0837; NACHT; 1.
KW Apoptosis; Repeat; Multigene family.
FT REPEAT 60 127
FT REPEAT 159 227
FT REPEAT 278 345
FT DOMAIN 508 802
FT CONFLICT 377 377 D -> G (IN REF. 3).
FT CONFLICT 403 403 L -> F (IN REF. 3).
FT CONFLICT 478 478 L -> I (IN REF. 3).
FT CONFLICT 540 540 N -> Y (IN REF. 3).
FT CONFLICT 862 862 K -> N (IN REF. 3).
FT CONFLICT 1079 1080 SD -> FN (IN REF. 3).

```

```
FT CONFLICT 1089 1089 R -> C (IN REF. 3).
FT CONFLICT 1115 1115 K -> E (IN REF. 3).
FT CONFLICT 1122 1122 T -> A (IN REF. 3).
FT CONFLICT 1136 1136 T -> E (IN REF. 3).
FT CONFLICT 1157 1157 S -> G (IN REF. 3).
FT CONFLICT 1167 1167 G -> R (IN REF. 3).
FT CONFLICT 1271 1271 F -> C (IN REF. 3).
SQ SEQUENCE 1447 AA; 164033 MW; 9EFC6A73BAE60A2 CRC64;

Query Match 18.3%; Score 142; DB 1; Length 1447;
Best Local Similarity 37.3%; Pred. No. 0.00012;
Matches 31; Conservative 13; Mismatches 35; Indels 4; Gaps 2;

QY 15 KDHRISTKWNPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFCKELEGWEPDDDP 74
DB 159 EEARLESFEDPFFY-AHGTSPRVLSAGVF---TGKRDVTQCFSCGGCLGNWEEGDDPW 214

QY 75 BEHKHSSGCAFLSVKQFEELT 97
DB 215 KEHAKWFKCEFLQSKKSPESIT 237

RESULT 13.
ID IAP GVCP STANDARD; PRT; 275 AA.
AC P41436;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Apoptosis inhibitor IAP.
GN Cydia pomonella granulosis virus (CpGV) (Cydia pomonella
OS Granulovirus).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
OX NCBI_TaxID=28289;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mexican 1;
RX MEDLINE=97380577; PubMed=8445726;
RA Crook N.E., Clem R.J., Miller L.K.;
RT "An apoptosis-inhibiting baculovirus gene with a zinc finger-like
RT motif.";
RN J. Virol. 67:2168-2174 (1993).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Mexican 1;
RX MEDLINE=97380577; PubMed=9237352;
RA Kang W., Crook N.E., Winstanley D., O'Reilly D.R.;
RT "Complete sequence and transposon mutagenesis of the BamHI J fragment
RT of Cydia pomonella granulosis virus.";
RN Virus Genes 14:131-136 (1997).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=Mexican 1;
RX MEDLINE=21445686; PubMed=11562546;
RA Luque T., Finch R., Crook N., O'Reilly D.R., Winstanley D.;
RT "The complete sequence of the Cydia pomonella granulovirus genome.";
RN J. Gen. Virol. 82:2531-2547 (2001).
CC -!- FUNCTION: ACTS BY BLOCKING CELLULAR APOPTOSIS RATHER THAN BY
CC PREVENTING VIRAL STIMULATION OF APOPTOSIS.
CC -!- SIMILARITY: Contains 2 BIR repeats.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U53466; AAB39098.1;
CC PIR; A45679; A45679.
DR
```

```
DR HSP; Q13490; IQBH.
DR InterPro; IPR001370; BIR.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00653; BIR; 2.
DR SMART; SM00238; BIR; 2.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01282; BIR_REPEAT 1; 2.
DR PROSITE; PS0143; BIR_REPEAT 2; 2.
DR PROSITE; PS00518; ZF_RING 1; FALSE_NEG.
DR PROSITE; PS00089; ZF_RING 2; 1.
KW Apoptosis; Zinc-finger; Repeat.
FT REPEAT 7 73 BIR 1.
FT REPEAT 108 175 BIR 2.
FT ZN_FING 228 263 RING-TYPE.
SQ SEQUENCE 275 AA; 31280 MW; 84605448869CAD60 CRC64;

Query Match 17.5%; Score 136.5; DB 1; Length 275;
Best Local Similarity 36.4%; Pred. No. 6.1e-05;
Matches 28; Conservative 14; Mismatches 24; Indels 11; Gaps 3;

QY 18 RISTKWNPFLEGGCACT---TPERMAEAGFIHCPTENEPDLAQCFCKELEGWEPDDDP 74
DB 111 RVKSFENWP-----RCMKQRPQOMADAGFFY---TGYGNTKCFYCDGLXWEPEDVFW 162

QY 75 BEHKHSSGCAFLSVKX 91
DB 163 EGVWFWFDRCAIVQLVK 179

RESULT 14
BIR7 HUMAN STANDARD; PRT; 298 AA.
ID Q96CA5; Q9BQV0; Q9H2A8; Q9HAP7;
AC Q96CA5; Q9BQV0; Q9H2A8; Q9HAP7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Baculoviral IAP repeat-containing protein 7 (kidney inhibitor of
DE apoptosis protein) (KIAP) (Melanoma inhibitor of apoptosis protein)
DE (ML-IAP) (Livin).
GN BIRC7 OR KIAP OR MLTAP OR LIVIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Petal kidney;
RX MEDLINE=21092523; PubMed=11162435;
RA Lin J.-H., Deng G., Huang Q., Morser J.;
RT "KIAP, a novel member of the inhibitor of apoptosis protein family.";
RN Biochem. Biophys. Res. Commun. 279:820-831 (2000).
[2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Melanoma;
RX MEDLINE=21223043; PubMed=11322947;
RA Ashhab Y., Alian A., Pollack A., Panet A., Yehuda D.B.;
RT "Two splicing variants of a new inhibitor of apoptosis gene with
RT different biological properties and tissue distribution pattern.";
RN FEBS Lett. 495:56-60 (2001).
[3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Kidney;
RX PubMed=11024045;
RA Kasof G.M., Gomes B.C.;
RT "Livin, a novel inhibitor of apoptosis protein family member.";
RN J. Biol. Chem. 276:3238-3246 (2001).
[4]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
```



[illegible]

```
FFCYGGLQSWKRGDDPWTEHAKWFP -> MFLPMDYPEAP
WRLHSSPPRCPRALQGGRRGNDPDRGSKGLKQTVGAL
NPNIHLWGIC (in isoform 3).
/FTid=VSP_002458.
Missing (in isoform 1).
/FTid=VSP_002459.
EE->AA: NO CHANGE IN SMAC INTERACTION AND
ANTI-APOPTOTIC ACTIVITY.
D->A: ABOLISHES INHIBITION OF CASPASES,
SMAC BINDING AND ANTI-APOPTOTIC ACTIVITY.
C->A: ABOLISHES INHIBITION OF CASPASES
AND ANTI-APOPTOTIC ACTIVITY.

Query Match 17.5%; Score 136.5; DB 1; Length 298;
Best Local Similarity 38.6%; Pred. No. 6.7e-05;
Matches 27; Conservative 9; Mismatches 29; Indels 5; Gaps 2;

QY 18 RISTFKNWPFLGGCACTPERMAEAGFIHCPTENEPDLAQCFCKELEGWEPDDDDPIEEH 77
Db 90 RLASFYDWPLT--AEVPELLAAAGFFH---TGHQDKVRFCFCYGGQLQSWKRGDDPWTEH 144
QY 78 KKHSSGCAFL 87
Db 145 AKWPPSCQFL 154

RESULT 15
IAP3_NPVOP
ID IAP3_NPVOP STANDARD; PRT; 268 AA.
AC P41437;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Apoptosis inhibitor 3 (IAP-3).
GN IAP3 OR IAP.
OS Orygia pseudotsugata multicapsid polyhedrosis virus (OpMPNV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_taxid=164623;
RN [1]_taxid=164623;
RP SEQUENCE FROM N.A.
RX MEDLINE=94187094; PubMed=8139034;
RA Birnbaum M.J., Clem R.J., Miller L.K.;
RT "An apoptosis-inhibiting gene from a nuclear polyhedrosis virus
RT encoding a polypeptide with Cys/His sequence motifs.";
RL J. Virol. 68:2521-2528(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97271300; PubMed=9126251;
RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RA Rohrmann G.F.;
RT "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
RT polyhedrosis virus genome.";
RL Virology 229:381-395(1997).
CC -!- FUNCTION: ACTS BY BLOCKING CELLULAR APOPTOSIS RATHER THAN BY
CC PREVENTING VIRAL STIMULATION OF APOPTOSIS.
CC -!- SIMILARITY: Contains 2 BIR repeats.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L22564; AB02610.1; -.
CC EMBL; U75930; AAC59034.1; -.
CC PIR; A53989; A53989.
CC HSSP; Q13490; 1QBH.
CC InterPro; IPR001370; BIR.
CC InterPro; IPR001841; Znf_ring.
```

```
DR Pfam; PF00653; BIR; 2.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00238; BIR; 2.
DR SMART; SM00184; BIR; 1.
DR PROSITE; PS01282; BIR_REPEAT_1; 2.
DR PROSITE; PS01443; BIR_REPEAT_2; 2.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS0089; ZF_RING_2; 1.
KW Apoptosis; Zinc-finger; Repeat.
FT REPEAT 18 84 BIR 1.
FT REPEAT 111 178 BIR 2.
FT ZN_FING 221 256 RING-TYPE
FT SEQUENCE 268 AA; 30076 MW; DF89175FDE85A708 CRC64;

Query Match 17.2%; Score 133.5; DB 1; Length 269;
Best Local Similarity 36.5%; Pred. No. 0.00011;
Matches 27; Conservative 12; Mismatches 30; Indels 5; Gaps 2;

QY 18 RISTFKNWPFLGGCACTPERMAEAGFIHCPTENEPDLAQCFCKELEGWEPDDDDPIEEH 77
Db 114 RLRTFAEWP--RGLQRPELAEAGFFY---TGQGDKTRCFCCDGLKDWEPDDAPWQOH 168
QY 78 KKHSSGCAFLSVKK 91
Db 169 ARWYDRCEYVLLVK 182
```

Search completed: August 11, 2004, 14:16:18  
Job time : 15 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 14:13:25 ; Search time 39 Seconds  
(without alignments)  
1148.810 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 778  
Sequence: 1 MGAPLPPAWQFLKDHRS.....EFETAKVRAIEQLAAMD 142

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirs.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	775	99.6	142	6	Q81009
2	718	92.3	142	6	Q9GUNS
3	556	71.5	121	11	Q923F7
4	486	62.5	142	13	Q9DDK0
5	456	58.6	157	13	Q804H7
6	444	57.1	135	13	Q90ZN0
7	389	50.0	142	13	Q90WU9
8	361	46.4	128	13	Q90WU8
9	356.5	45.8	160	13	Q8UG75
10	356.5	45.8	160	13	Q8JGN5
11	351.5	45.2	160	13	Q8AXI6
12	246.5	31.7	153	5	Q9VSM2
13	160	20.6	524	11	Q8NX0
14	160	20.6	865	11	Q8BR72
15	160	20.6	4845	11	Q88738
16	158	20.3	1151	5	Q8MS47

17	158	20.3	4876	5	Q8MUM9
18	158	20.3	4904	5	Q9VH01
19	157	20.2	308	5	Q18727
20	155.5	20.0	281	12	Q9YNL8
21	149	19.2	288	12	Q9IF18
22	148	19.0	288	12	Q9E232
23	146.5	18.8	405	13	Q8UWH2
24	146.5	18.8	415	13	Q7SXU1
25	145.5	18.7	155	5	Q22837
26	145	18.6	346	5	Q8IS31
27	145	18.6	1403	11	Q8CH68
28	145	18.6	1403	11	Q8CH64
29	143	18.4	346	5	Q968F8
30	143	18.4	597	11	Q9R015
31	143	18.4	1402	11	Q8CH65
32	143	18.4	1402	11	Q8CGT4
33	143	18.4	1402	11	Q8CGT3
34	143	18.4	1403	11	Q8CH70
35	143	18.4	1403	11	Q8CGT2
36	142	18.3	283	12	Q80LK3
37	142	18.3	1271	11	Q8CH71
38	142	18.3	1432	11	Q8CGS9
39	142	18.3	1446	11	Q8CG17
40	142	18.3	1447	11	Q8CH66
41	142	18.3	1447	11	Q8CGT1
42	142	18.3	1447	11	Q8CGT0
43	142	18.3	1447	11	Q8CGS8
44	142	18.3	1447	11	Q8CGS7
45	140.5	18.1	249	5	Q8SVK6

#### ALIGNMENTS

##### RESULT 1

Q81009 PRELIMINARY; PRT; 142 AA.  
ID Q81009  
AC Q81009;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Survivin.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_taxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Inoue C., Yamazaki J., Kano R., Hasegawa A.;  
RT "Canis familiaris mRNA for survivin-protein, complete cds."  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB0951108; BAC22748.2;  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.  
DR GO; GO:0006916; F:anti-apoptosis; IEA.  
DR InterPro; IPR001370; BIR.  
DR Pfam; PF00653; BIR.  
DR SMART; SM00238; BIR.  
DR PROSITE; PS50143; BIR. REPEAT 2; 1.  
SQ SEQUENCE 142 AA; 16375 MW; E960DC9DE6D36792 CRC64;

Query Match 99.6%; Score 775; DB 6; Length 142;

Best Local Similarity 99.3%; Pred. No. 1.5e-64;

Matches 141; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGAPLPPAWQFLKDHRSIFKWPFLGGCACTPERMAEAGTHCPTENEPDLAQCFPC 60

Db 1 MGAPLPPAWQFLKDHRSIFKWPFLGGCACTPERMAEAGTHCPTENEPDLAQCFPC 60

Qy 61 PKELEGWPEDDPTEEHKHSKGCAFLSVKKQFELTIGELKLDREAFKAKIATNNK 120

Db 61 PKELEGWPEDDPTEEHKHSKGCAFLSVKKQFELTIGELKLDREAFKAKIATNNK 120

```

QY 121 KKEFEETAKVRRATEQLAAMD 142
DB 121 KKEFEETAKVRRATEQLAAMD 142

RESULT 2
Q9GLNS
ID Q9GLN5 PRELIMINARY; PRT; 142 AA.
AC Q9GLN5;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Apoptosis inhibitor survivin.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21328903; PubMed=11435714;
RA Cirera S., Fredholm M.;
RT "Isolation and mapping the pig homologs survivin (BIRC5) and effector
RL cell protease receptor 1 (APR1) genes.";
DR EMBL; AF195781; AAG17540.1; -.
DR HSSP; O15392; IE31.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO; GO:0006916; P:anti-apoptosis; IEA.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 1.
DR SMART; SM00238; BIR; 1.
DR PROSITE; PS0143; BIR REPEAT 2; 1.
SQ SEQUENCE 142 AA; 16297 MW; 56BA753B651F4F2C CRC64;

Query Match 92.3%; Score 718; DB 6; Length 142;
Best Local Similarity 92.3%; Pred. No. 3e-59; Indels 0; Gaps 0;
Matches 131; Conservative 5; Mismatches 6;

QY 1 MGAFTLPPAWQPFLLKDHRISTFKNWPFLGCACTPERMAEAGFIHCTPTNEPDLAQCF 60
DB 1 MSAPSLPPAWQLKDHRISTFKNWPFLGCACTPERMAAAGFIHCTPTNEPDLAQCF 60

QY 61 FKELEGWEPDDPIEHHKSHSGCAFLSVKQFELTLGFLKLDREKAKNIKAKETNNK 120
DB 61 FKELEGWEPDDPIEHHKSHSGCAFLSVKQFELTLGFLKLDREKAKNIKAKETNNK 120

QY 121 KKEFEETAKVRRATEQLAAMD 142
DB 121 KKEFEETAKVRRATEQLAAMD 142

RESULT 3
Q923F7
ID Q923F7 PRELIMINARY; PRT; 121 AA.
AC Q923F7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Baculoviral IAP repeat-containing 5.
GN BIRC5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004702; AA04702.1; -.
DR MGP; MGI:1203517; Birc5.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0008189; F:apoptosis inhibitor activity; IMP.

```

```

DR GO; GO:0006916; P:anti-apoptosis; IMP.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 1.
DR SMART; SM00238; BIR; 1.
DR PROSITE; PS0143; BIR REPEAT 2; 1.
SQ SEQUENCE 121 AA; 14154 MW; 1E0EC7E01BA65585 CRC64;

Query Match 71.5%; Score 556; DB 11; Length 121;
Best Local Similarity 81.7%; Pred. No. 3.1e-44;
Matches 98; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 1 MGAFTLPPAWQPFLLKDHRISTFKNWPFLGCACTPERMAEAGFIHCTPTNEPDLAQCF 60
DB 1 MGAFTLPPAWQPFLLKDHRISTFKNWPFLGCACTPERMAEAGFIHCTPTNEPDLAQCF 60

QY 61 FKELEGWEPDDPIEHHKSHSGCAFLSVKQFELTLGFLKLDREKAKNIKAKETNNK 120
DB 61 FKELEGWEPDDPIEHHKSHSGCAFLSVKQFELTLGFLKLDREKAKNIKAKETNNK 120

RESULT 4
Q9DDK0
ID Q9DDK0 PRELIMINARY; PRT; 142 AA.
AC Q9DDK0;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Survivin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX Langer J.S., Johnson A.L., Bridgham J.T.;
RL "Molecular cloning and characterization of chicken survivin.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF322051; AAG42494.1; -.
DR HSSP; O15392; IE31.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO; GO:0006916; P:anti-apoptosis; IEA.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 1.
DR SMART; SM00238; BIR; 1.
DR PROSITE; PS0143; BIR REPEAT 2; 1.
SQ SEQUENCE 142 AA; 15955 MW; 1323CCBD5EE2F551 CRC64;

Query Match 62.5%; Score 486; DB 13; Length 142;
Best Local Similarity 61.8%; Pred. No. 1.2e-37;
Matches 84; Conservative 19; Mismatches 33; Indels 0; Gaps 0;

QY 3 APTLPPAWQPFLLKDHRISTFKNWPFLGCACTPERMAEAGFIHCTPTNEPDLAQCF 62
DB 5 AEMLPKWLVLVSTRAATFRNWPFTGCACTPERMAAAGFVHCPSENPVXCFFCLK 64

QY 63 ELEGWEPDDPIEHHKSHSGCAFLSVKQFELTLGFLKLDREKAKNIKAKETNNK 122
DB 65 ELEGWEPDDPIEHHKSHSGCAFLSVKQFELTLGFLKLDREKAKNIKAKETNNK 124

QY 123 EFEETAKVRRATEQL 138
DB 125 DIEDVAKGVRAIENM 140

RESULT 5
Q804H7
ID Q804H7 PRELIMINARY; PRT; 157 AA.
AC Q804H7;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

```

[illegible]

OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2037392; PubMed=10917738;  
 RA Inohara N., Nunez G.;  
 RT "Genes with homology to mammalian apoptosis regulators identified in  
 zebrafish";  
 RL Cell Death Differ. 7:509-510(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Inohara N., Nunez G.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY057058; AAL18251.1; -  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0008189; P:apoptosis inhibitor activity; IEA.  
 DR GO; GO:0006916; P:anti-apoptosis; IEA.  
 DR InterPro; IPR001370; BIR.  
 DR Pfam; PF00653; BIR; 1.  
 DR SMART; SM00238; BIR; 1.  
 DR PROSITE; PS50143; BIR REPEAT 2; 1.  
 SQ SEQUENCE 138 AA; 15294 MW; A8496849082DDC66 CRC64;  
 Query Match 46.4%; Score 361; DB 13; Length 128;  
 Best Local Similarity 52.4%; Pred. No. 4.8e-25;  
 Matches 65; Conservative 19; Mismatches 36; Indels 4; Gaps 1;  
 QY 16 DHRISTFKNWPFLGCACTPERMAEAGFIHCPTENEPDLAQCFFCFKELEGWEPDDDDPIE 75  
 DB 5 EKRLCTSEWFRDCCQCTPLMAKAGFVHCPSNEPDVACCPCLKELEGWEPDDNPS 64  
 QY 76 EHKHSSGCAFLSVKKQFEELTLGEFLKLDREARAKNIATK-ETNNKKKEFEETAKKV 131  
 DB 65 EHAKRSPNCAFLHMSKTFDELTAIEYHLEQERLIRVIKMGHRKIAYPRDEVATSKNL 124  
 QY 132 RRAI 135  
 DB 125 RALI 128  
 RESULT 9  
 Q8JG75 PRELIMINARY; PRT; 160 AA.  
 AC Q8JG75;  
 DT 01-OCT-2002 (TRENBLrel. 22, Created)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Survivin.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bolton M.A., Lan W., Powers S.E., McClelland M.L., Kuang J.,  
 RA Stukenberg P.T.;  
 RT "Aurora B kinase exists in a complex with survivin and INCENP and its  
 RT kinase activity is stimulated by survivin binding and  
 RT phosphorylation";  
 RL Mol. Biol. Cell 0:0-0(2002).  
 DR EMBL; AY115553; AAM76714.1; -  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0008189; P:apoptosis inhibitor activity; IEA.  
 DR GO; GO:0006916; P:anti-apoptosis; IEA.  
 DR InterPro; IPR001370; BIR.  
 DR Pfam; PF00653; BIR; 1.  
 DR SMART; SM00238; BIR; 1.  
 DR PROSITE; PS50143; BIR REPEAT 2; 1.  
 SQ SEQUENCE 160 AA; 18677 MW; B7C32E02119DC61B CRC64;  
 Query Match 45.8%; Score 356.5; DB 13; Length 160;  
 Best Local Similarity 54.1%; Pred. No. 1.6e-25;  
 Matches 66; Conservative 17; Mismatches 38; Indels 1; Gaps 1;  
 QY 16 DHRISTFKNWPFLGCACTPERMAEAGFIHCPTENEPDLAQCFFCFKELEGWEPDDDDPIE 75  
 DB 25 DARLATFADWPFTECKCTPESMAKAGFVHCPTENEPDVACCFCLKELEGWEPDDDPWT 84  
 QY 76 EHKHSSGCAFLSVKKQFEELTLGEFLKLDREARAKNIATK-ETNNKKKEFEETAKKV 134  
 DB 85 EHSKESANGCFLSLTKVCVNDLTMEGFLRLEGDRIKSFYRKESTVVLQYVEEMTAATKRL 144  
 QY 135 IE 136  
 DB 145 LE 146  
 RESULT 11  
 Q8AXI6 PRELIMINARY; PRT; 160 AA.  
 ID Q8AXI6  
 AC Q8AXI6;  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Survivin.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.

Matches 66; Conservative 17; Mismatches 38; Indels 1; Gaps 1;  
 QY 16 DHRISTFKNWPFLGCACTPERMAEAGFIHCPTENEPDLAQCFFCFKELEGWEPDDDDPIE 75  
 DB 25 DARLATFADWPFTECKCTPESMAKAGFVHCPTENEPDVACCFCLKELEGWEPDDDPWT 84  
 QY 76 EHKHSSGCAFLSVKKQFEELTLGEFLKLDREARAKNIATK-ETNNKKKEFEETAKKV 134  
 DB 85 EHSKESANGCFLSLTKVCVNDLTMEGFLRLEGDRIKSFYRKESTVVLQYVEEMTAATKRL 144  
 QY 135 IE 136  
 DB 145 LE 146  
 RESULT 10  
 Q8JGN5 PRELIMINARY; PRT; 160 AA.  
 ID Q8JGN5  
 AC Q8JGN5;  
 DT 01-OCT-2002 (TRENBLrel. 22, Created)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Survivin/XBIR.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Losada A., Hirano M., Hirano T.;  
 RT "Aurora B and polo kinases cooperate to establish metaphase chromosome  
 RT structure with bi-oriented sister chromatids";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY100639; AAM44085.1; -  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0008189; P:apoptosis inhibitor activity; IEA.  
 DR GO; GO:0006916; P:anti-apoptosis; IEA.  
 DR InterPro; IPR001370; BIR.  
 DR Pfam; PF00653; BIR; 1.  
 DR SMART; SM00238; BIR; 1.  
 DR PROSITE; PS50143; BIR REPEAT 2; 1.  
 SQ SEQUENCE 160 AA; 18686 MW; 9AC32E02119DC609 CRC64;  
 Query Match 45.8%; Score 356.5; DB 13; Length 160;  
 Best Local Similarity 54.1%; Pred. No. 1.6e-25;  
 Matches 66; Conservative 17; Mismatches 38; Indels 1; Gaps 1;  
 QY 16 DHRISTFKNWPFLGCACTPERMAEAGFIHCPTENEPDLAQCFFCFKELEGWEPDDDDPIE 75  
 DB 25 DARLATFADWPFTECKCTPESMAKAGFVHCPTENEPDVACCFCLKELEGWEPDDDPWT 84  
 QY 76 EHKHSSGCAFLSVKKQFEELTLGEFLKLDREARAKNIATK-ETNNKKKEFEETAKKV 134  
 DB 85 EHSKESANGCFLSLTKVCVNDLTMEGFLRLEGDRIKSFYRKESTVVLQYVEEMTAATKRL 144  
 QY 135 IE 136  
 DB 145 LE 146  
 RESULT 11  
 Q8AXI6 PRELIMINARY; PRT; 160 AA.  
 ID Q8AXI6  
 AC Q8AXI6;  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Survivin.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.

```

OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22341152; PubMed=12454937;
RA Murphy C.R., Sabel J.L., Sandler A.D., Dagle J.M.;
RT "Survivin mRNA is down-regulated during early Xenopus laevis
   embryogenesis.";
RL Dev. Dyn. 225:597-601(2002).
DR EMBL; AF442492; AAN76690.1; -.
DR GO; GO:0003622; C:intracellular; IEA.
DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO; GO:0008916; P:anti-apoptosis; IEA.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 1.
DR SMART; SM00238; BIR; 1.
DR PROSITE; PS0143; BIR REPEAT 2; 1.
SQ SEQUENCE 160 AA, 18812 MW, 57538991.06190F4B CRC64;

Query Match 45.2%; Score 351.5; DB 13; Length 160;
Best Local Similarity 54.2%; Pred. No. 4.6e-25;
Matches 65; Conservative 16; Mismatches 38; Indels 1; Gaps 1;

QY 18 RISTKNWPFLEGCACPTPMAAGFIHCPTENEPDLAQCFCKELEGWEPDDDPPIBEH 77
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
27 RLATFADWFTENCXCTPENWAKAGVHCPTENEPDVACCFCKELEGWEPDDDPWNEH 86
QY 78 KHSSCCAFSLVKQFEETLGFELKLDREAKNKIAK-ETNKKKPEPETAKKVRAIE 136
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
87 SKRSVNCGLSLTKVNDLTMEGFLLEGDRISFYRKFTSTVLQVSEEMTAATKRLLE 146

RESULT 12
Q9VEM2 PRELIMINARY; PRT; 153 AA.
AC Q9VEM2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cg12265 protein (B555472p).
GN Cg12265.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Fendell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck J., Brockstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry J., Morris J., Moshrefi A.,

```

```

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,
RA Svirskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Achavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Lao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleab J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003716; AAP55399.1; -.
DR EMBL; AY071487; AAL49109.1; -.
DR HSPF; C15392; 1531.
DR Flybase; FBgn0038489; CG12265.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO; GO:0006916; P:anti-apoptosis; IEA.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 1.
DR SMART; SM00238; BIR; 1.
DR PROSITE; PS0143; BIR REPEAT 2; 1.
SQ SEQUENCE 153 AA, 17455 MW, EE250E3B669EC359 CRC64;

Query Match 31.7%; Score 246.5; DB 5; Length 153;
Best Local Similarity 40.3%; Pred. No. 2.6e-15;
Matches 48; Conservative 20; Mismatches 50; Indels 1; Gaps 1;

QY 14 LKHRISTFRNWPFLGCACPTPMAAGFIHCPTENEPDLAQCFCKELEGWEPDDDP 73
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
27 LEQHRVSYKSWPPFPETASCSISKVABAGFYWTGKRENDTATCFVCGKTLDGWEPEDDP 86
QY 74 IEHKKHSSGCCAFSLVKQFEETLGFELKLDREAKNKIAK-ETNKKKPEF-BETAKKV 131
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
87 WKEHVKAPOCFEAKLSCPERNLTVSQFLGLTVGVKSGIEKTKAFKFSFVRENEKRL 145

RESULT 13
Q8BNX0 PRELIMINARY; PRT; 524 AA.
AC Q8BNX0
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Baculoviral IAP repeat-containing 6 (Fragment).
DE A430040A191K.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta, and Vein;
RX MEDLINE=22354683; PubMed=12466851;
RA THE FANTOM Consortium,
RA THE RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
   60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RL EMBL; AK079995; BAC37801.1; -.

```

```

DR PIR; PT0546; PT0698.
DR PIR; PT0675; PT0675.
DR MGD; MGI:2444861; A430040A19R1k.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO; GO:0006916; P:anti-apoptosis; IEA.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 1.
DR SMART; SM00238; BIR; 1.
DR PROSITE; PS50143; BIR_REPEAT_2; 1.
FT NON_TER 524 524
SQ SEQUENCE 524 AA; 56194 MW; 6DA7DF987347D6FD CRC64;

Query Match      20.6%; Score 160; DB 11; Length 865;
Best Local Similarity 42.3%; Pred. No. 1.9e-06;
Matches 30; Conservative 10; Mismatches 31; Indels 0; Gaps 0;

QY 17 HRISTFKNWPFLGCACTPERMAEAGFIHCTENEPDLAQCFCKELEGWEPDDPIEE 76
Db 263 NRRTFTTSWPHVGYRWAQPDMAQAGFYHQPASSGDDDRAMCFTCSVCLVCWEPTDEPSE 322
QY 77 HKHSSGCAFL 87
Db 323 HERHSPNCPFV 333

RESULT 14
Q8BR72 PRELIMINARY; PRT; 865 AA.
AC Q8BR72;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Baculoviral IAP repeat-containing 6 (Fragment).
GN A430040A19R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK045446; BAC32373.1; -.
DR PIR; PT0546; PT0698.
DR MGD; MGI:2444861; A430040A19R1k.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO; GO:0006916; P:anti-apoptosis; IEA.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 1.
DR SMART; SM00238; BIR; 1.
DR PROSITE; PS50143; BIR_REPEAT_2; 1.
FT NON_TER 865 865
SQ SEQUENCE 865 AA; 93786 MW; EB0EF6FB27ACB6D CRC64;

Query Match      20.6%; Score 160; DB 11; Length 865;
Best Local Similarity 42.3%; Pred. No. 1.9e-06;
Matches 30; Conservative 10; Mismatches 31; Indels 0; Gaps 0;

QY 17 HRISTFKNWPFLGCACTPERMAEAGFIHCTENEPDLAQCFCKELEGWEPDDPIEE 76
Db 263 NRRTFTTSWPHVGYRWAQPDMAQAGFYHQPASSGDDDRAMCFTCSVCLVCWEPTDEPSE 322
QY 77 HKHSSGCAFL 87
Db 323 HERHSPNCPFV 333

```

```

RESULT 15
Q8BR73 PRELIMINARY; PRT; 4845 AA.
AC Q8BR73;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ubiquitin-conjugating enzyme E2 (EC 6.3.2.19) (Ubiquitin-protein
DE ligase) (Ubiquitin carrier protein).
GN BIRC6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98292517; PubMed=9628897;
RA Hauser H.P.; Bardroff M.; Pyrowolakis G.; Jentsch S.;
RT "A giant ubiquitin-conjugating enzyme related to IAP apoptosis
RT inhibitors.";
RL J. Cell Biol. 141:1415-1422(1998).
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC -!- PATHWAY: UBIQUITIN CONJUGATION; SECOND STEP.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL; Y1267; CAA76720.1; -.
DR PIR; T31067; T31067.
DR HSSP; Q13490; IQBH.
DR MGD; MGI:1276108; Birc6.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006916; P:anti-apoptosis; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR001370; BIR.
DR InterPro; IPR000608; UBQ_conjugat.
DR Pfam; PF00653; BIR; 1.
DR Pfam; PF00179; UQ_con; 1.
DR ProDom; PD000461; UBQ_conjugat; 1.
DR SMART; SM00238; BIR; 1.
DR SMART; SM00212; UBCC; 1.
DR PROSITE; PS50143; BIR_REPEAT_2; 1.
DR PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.
DR KW
DR Ligase; Ubl conjugation pathway.
DR SEQUENCE 4845 AA; 528411 MW; 35C585F73B8CF84C CRC64;

Query Match      20.6%; Score 160; DB 11; Length 4845;
Best Local Similarity 42.3%; Pred. No. 1.2e-05;
Matches 30; Conservative 10; Mismatches 31; Indels 0; Gaps 0;

QY 17 HRISTFKNWPFLGCACTPERMAEAGFIHCTENEPDLAQCFCKELEGWEPDDPIEE 76
Db 263 NRRTFTTSWPHVGYRWAQPDMAQAGFYHQPASSGDDDRAMCFTCSVCLVCWEPTDEPSE 322
QY 77 HKHSSGCAFL 87
Db 323 HERHSPNCPFV 333

```

Search completed: August 11, 2004, 14:17:08  
Job time : 40 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 14:10:05 ; Search time 53 Seconds

(without alignments)  
757.014 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 778

Sequence: 1 MGAPLPPAWQFLKDHRS.....EFETAKVRAIEQLAAMD 142

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*\*

- 1: Geneseq1980s:\*\*
- 2: Geneseq1990s:\*\*
- 3: Geneseq2000s:\*\*
- 4: Geneseq2001s:\*\*
- 5: Geneseq2002s:\*\*
- 6: Geneseq2003as:\*\*
- 7: Geneseq2003bs:\*\*
- 8: Geneseq2004s:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	778	100.0	142	2	AAW61359 Survivin
2	778	100.0	142	2	Aay43080 Human Sur
3	778	100.0	142	4	Ag65360 Human Her
4	778	100.0	142	5	Aau74571 Human sur
5	778	100.0	142	5	Aao14948 Survivin-
6	778	100.0	142	6	Abp72162 Human sur
7	778	100.0	142	7	Abp61582 Human sur
8	778	100.0	142	4	Aau29999 Novel hum
9	776	99.7	142	5	Aau74572 Human mut
10	772	99.2	142	5	Aau74574 Human mut
11	772	99.2	142	5	Aau74577 Human mut
12	768	98.7	142	5	Aau74575 Human mut
13	764	98.2	142	5	Aau74573 Human mut
14	762	98.0	141	6	Abu56477 Lung canc
15	762	97.9	142	5	Aau74576 Human mut
16	756.5	97.2	165	6	Ada09936 Human sur
17	750	96.4	142	5	Aau74580 Human mut
18	744	95.6	142	5	Aau74578 Human mut
19	742	95.4	142	5	Aau74581 Human mut
20	736	94.6	142	5	Aau74579 Human mut
21	699	89.8	125	5	Aau74582 Human del
22	660	84.8	140	2	AAW19749 Mouse inh
23	660	84.8	140	5	Aau74583 Mouse inh
24	642	82.5	116	3	AGG02311 Human sec
25	436	56.0	75	7	ADB61824 Human inh

26	354	45.5	83	7	ADB61823	Human inh
27	246.5	31.7	153	4	ABB61397	Drosophil
28	246.5	31.7	153	7	ADC06777	Fruit fly
29	217	27.9	155	4	AAU30231	Novel hum
30	160	20.6	75	7	ADB61822	Human inh
31	160	20.6	4829	4	AA897833	Human apo
32	158	20.3	4904	4	ABB62249	Drosophil
33	157	20.2	308	2	AAW19750	C. elegan
34	155	19.9	1232	2	AAAR98217	Neuronal
35	155	19.9	1295	2	AAAY09540	Human apo
36	155	19.9	1295	2	AAAY14080	Gonadotro
37	155	19.9	1403	2	AAW20033	Neuronal
38	155	19.9	1403	2	AAW20032	Neuronal
39	155	19.9	1403	2	AAAY09539	Human apo
40	155	19.9	1403	2	AAAY14079	Gonadotro
41	155	19.9	1403	3	AAAY88053	Human NAI
42	155	19.9	1403	6	ABU07400	Protein d
43	155	19.9	1403	6	ABP72166	Human inh
44	145.5	18.7	155	2	AAW19751	C. elegan
45	143	18.4	346	5	ABB78046	Amino aci

ALIGNMENTS

RESULT 1  
AAW61359  
ID AAW61359 standard; protein; 142 AA.

XX AAW61359;

DT 25-SEP-1998 (first entry)

DE Survivin protein.

survivin; apoptosis; cellular apoptosis; transplantation;

motor neuron degenerative disease; HIV infection; immunosuppression;

gastrointestinal perturbations; cardiovascular disorder.

OS Homo sapiens.

XX WO9822589-A2.

XX 28-MAY-1998.

XX 20-NOV-1997; 97WO-US021880.

XX 20-NOV-1996; 96US-0031435P.

XX 20-NOV-1997; 97US-00975080.

(UYVA ) UNIV YALE.

XX Altieri DC;

XX WPI; 1998-312475/27.

XX N-PSDB; AAV27941.

Modulating apoptosis by controlling the Survivin gene - useful for

treating transplant rejection, degenerative disorders and tumours.

PS Disclosure; Fig 10; 108pp; English.

XX The survivin gene can be used to control apoptosis through modification  
of the gene. Survivin peptides can be used to inhibit cellular apoptosis,  
e.g. for enhancing the viability of organs and tissues prior to their  
transplantation, for preserving the growth of cells in culture or for  
treating conditions involving abnormal apoptosis, e.g. degenerative  
diseases such as motor neuron degenerative diseases, HIV infection,  
dermatological effects of ageing, disorders and diseases such as  
immunosuppression, gastrointestinal perturbations, cardiovascular  
disorders, apoptosis related to reperfusion damage, rejection of tissue  
transplantation and Alzheimer's disease. Agents which block Survivin  
activity can be used to treat e.g. tumours

XX SQ Sequence 142 AA;  
 Query Match 100.0%; Score 778; DB 2; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-78;  
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAPTLPPAWQPFKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60  
 DB 1 MGAPTLPPAWQPFKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60  
 QY 61 FKELEGWEPDDPIEBHKHSSGCAFLSVKKQFELTLGFLKDRERAKNKIAKETNNK 120  
 DB 61 FKELEGWEPDDPIEBHKHSSGCAFLSVKKQFELTLGFLKDRERAKNKIAKETNNK 120  
 QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 DB 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 2  
 AAY49080  
 ID AAY49080 standard; protein; 142 AA.  
 XX  
 AC AAY49080;  
 XX  
 DT 17-JAN-2000 (first entry)  
 XX  
 DE Human Survivin amino acid sequence.  
 XX  
 KW Survivin; IAP; apoptosis inhibitor; tumour cell; mitotic spindle;  
 KW effector cell protease receptor-1; EPR-1; embryonic tissue; tubulin;  
 KW abnormal cell growth; tumorigenesis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9950440-A2.  
 XX  
 PD 07-OCT-1999.  
 XX  
 PF 01-APR-1999; 99WO-US007205.  
 XX  
 PR 01-APR-1998; 98US-0080288P.  
 XX  
 PA (UYVA ) UNIV YALE.  
 XX  
 PI Altieri DC;  
 XX  
 DR WPI; 1999-591329/50.  
 DR N-PSDB; AAZ31037.  
 XX  
 PT Identifying agents that modulate Survivin interactions.  
 XX  
 PS Example; Fig 8A; 56pp; English.  
 XX  
 CC This is the amino acid sequence of the human survivin protein. Survivin  
 CC is a 142 amino acid protein of approximately 16 kD. Survivin is a member  
 CC of the IAP family of apoptosis inhibitors, and the gene is located on  
 CC chromosome 17q25. The nucleic acid sequence of Survivin is related to the  
 CC Effector Cell Protease Receptor-1 (EPR-1). Survivin is expressed in  
 CC tumour cells and embryonic tissue. The sequence is used in the invention  
 CC which involves identifying an agent that modulates interactions between  
 CC Survivin and tubulin. Agents that modulate interaction between Survivin  
 CC and tubulin or mitotic spindles may be administered to a mammal to  
 CC modulate biological or pathological processes mediated by Survivin, for  
 CC example Survivin-mediated inhibition of cellular apoptosis. Thus an  
 CC inhibiting agent may be used to block abnormal cell growth, for example  
 CC during tumorigenesis. An agent that increases interaction may be used to  
 CC extend cell growth in culture  
 XX  
 SQ Sequence 142 AA;  
 Query Match 100.0%; Score 778; DB 2; Length 142;

Best Local Similarity 100.0%; Pred. No. 7.9e-78;  
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAPTLPPAWQPFKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60  
 DB 1 MGAPTLPPAWQPFKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60  
 QY 61 FKELEGWEPDDPIEBHKHSSGCAFLSVKKQFELTLGFLKDRERAKNKIAKETNNK 120  
 DB 61 FKELEGWEPDDPIEBHKHSSGCAFLSVKKQFELTLGFLKDRERAKNKIAKETNNK 120  
 QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 DB 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 3  
 AAG65360  
 ID AAG65360 standard; protein; 142 AA.  
 XX  
 AC AAG65360;  
 XX  
 DT 30-NOV-2001 (first entry)  
 XX  
 DE Human Her-3 polypeptide sequence.  
 XX  
 KW Her-3; epidermal growth factor; EGF; receptor/tyrosine kinase; human;  
 KW antiinflammatory; cytostatic; antibacterial; antisense.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6277640-B1.  
 XX  
 PD 21-AUG-2001.  
 XX  
 PF 31-JUL-2000; 2000US-00630706.  
 XX  
 PR 31-JUL-2000; 2000US-00630706.  
 XX  
 PA (ISIS-) ISIS PHARM INC.  
 XX  
 PI Bennett CP, Cowser LM;  
 XX  
 DR WPI; 2001-535134/59.  
 DR N-PSDB; AAH47531.  
 XX  
 PT Antisense compounds capable of modulating expression of human Her-3,  
 PT member of epidermal growth factor family of receptor/tyrosine kinases,  
 PT useful for preventing or delaying infection, inflammation or tumor  
 PT formation.  
 XX  
 PS Example 15; Col 57-70; 49pp; English.  
 XX  
 CC The invention provides antisense compounds capable of inhibiting the  
 CC expression of human Her-3, a member of epidermal growth factor (EGF)  
 CC family of receptor/tyrosine kinases. The antisense oligonucleotides are  
 CC useful for inhibiting the expression of Her-3 in cells or tissues. They  
 CC are commonly used as research reagents and in diagnostics for example, to  
 CC elucidate the function of particular genes. The antisense compounds are  
 CC also useful for distinguishing between functions of various members of a  
 CC biological pathway and for research use. They are also utilized for  
 CC diagnostics, therapeutics, prophylaxis and in kits. They are useful  
 CC prophylactically, e.g. to prevent or delay infection, inflammation or  
 CC tumor formation. The present sequence represents the human Her-3 sequence  
 CC (GenBank Accession No. U75285).  
 XX  
 SQ Sequence 142 AA;  
 Query Match 100.0%; Score 778; DB 4; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-78;  
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAPTLPPAWQPFKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60

```

Db      1 MGAPLPPAWQPFVKDHRISTFKNWPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60
        |||||||
QY      61 FKELEGWEPDDPIEHHKHSSGCAFLSVKKQFEEELTGEFLKLDREAKNKIAKETNNK 120
        |||||||
Db      61 FKELEGWEPDDPIEHHKHSSGCAFLSVKKQFEEELTGEFLKLDREAKNKIAKETNNK 120
        |||||||
QY      121 KKEFEETAKKVRRAIEQLAAMD 142
        |||||||
Db      121 KKEFEETAKKVRRAIEQLAAMD 142
        |||||||

RESULT 4
AAU74571
ID      AAU74571 standard; protein; 142 AA.
XX
AC      AAU74571;
XX
XX
DT      08-MAY-2002 (first entry)
XX
DE      Human survivin polypeptide.
XX
KW      Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;
KW      tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;
KW      dimerisation; ovary; breast; pancreas; central nervous system; blood;
KW      lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;
KW      thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;
KW      cell proliferative disorder; acquired immunodeficiency syndrome;
KW      neurodegenerative disease; ischaemic injury; toxin-induced liver disease;
KW      myelodysplastic syndrome; protein coordinate data.
XX
OS      Homo sapiens.
XX
XX
PN      WO200202622-A2.
XX
XX
PD      10-JAN-2002.
XX
XX
PF      29-JUN-2001; 2001WO-US020872.
XX
XX
PR      29-JUN-2000; 2000US-00608352.
XX
XX
PA      (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
XX
PI      Noel JP, Verdacia M, Hunter T, Huang H;
XX
XX
DR      WPI; 2002-171638/22.
XX
XX
PT      Novel isolated crystalline survivin polypeptide, useful to screen and
PT      design compounds that bind to or interact with inhibitor of apoptosis
PT      protein and protein family members, and for design of novel therapeutics.
XX
XX
PS      Claim 42; Fig 2; 102pp; English.
XX
XX
CC      The invention relates to an isolated crystalline human survivin
CC      polypeptide, an inhibitor or apoptosis protein (IAP). The polypeptide is
CC      useful for determining the ability of a survivin binding agent to
CC      modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc
CC      chelation activity and for identifying an agent such as a peptide or
CC      peptidomimetic which inhibits dimerisation of survivin. The ability of an
CC      agent to modulate dimerisation can be determined through detection of a
CC      change in apoptosis in a target cell expressing survivin. Binding agents
CC      are useful for increasing apoptosis in a cell derived from a tissue
CC      selected from ovary, breast, pancreas, lymph node, skin, blood, lung,
CC      brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous
CC      system, prostate, colon, rectum, cervix or endometrium, with a cell
CC      proliferative disorder such as cancer. The molecules can also be used in
CC      treatment or prevention of apoptosis which occurs as a part of AIDS,
CC      neurodegenerative diseases, ischaemic injury, toxin-induced liver disease
CC      and myelodysplastic syndromes. This sequence represents the human
CC      survivin polypeptide
XX
XX
SQ      Sequence 142 AA;

```

```

Query Match      100.0%; Score 778; DB 5; Length 142;
Best Local Similarity 100.0%; Pred. No. 7.9e-78;
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MGAPLPPAWQPFVKDHRISTFKNWPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60
        |||||||
Db      1 MGAPLPPAWQPFVKDHRISTFKNWPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60
        |||||||
QY      61 FKELEGWEPDDPIEHHKHSSGCAFLSVKKQFEEELTGEFLKLDREAKNKIAKETNNK 120
        |||||||
Db      61 FKELEGWEPDDPIEHHKHSSGCAFLSVKKQFEEELTGEFLKLDREAKNKIAKETNNK 120
        |||||||
QY      121 KKEFEETAKKVRRAIEQLAAMD 142
        |||||||
Db      121 KKEFEETAKKVRRAIEQLAAMD 142
        |||||||

RESULT 5
AAO14948
ID      AAO14948 standard; protein; 142 AA.
XX
XX
AC      AAO14948;
XX
XX
DT      06-AUG-2002 (first entry)
XX
DE      Survivin-like protein 6.
XX
KW      Survivin-like protein; diagnosis; screening; cancer;
KW      apoptosis abnormality; gene therapy.
XX
XX
OS      Unidentified.
XX
XX
PN      WO200233071-A1.
XX
XX
PD      25-APR-2002.
XX
XX
PF      16-OCT-2001; 2001WO-JP009071.
XX
XX
PR      17-OCT-2000; 2000JP-00316721.
XX
XX
PR      20-DEC-2000; 2000JP-00386809.
XX
XX
PA      (TAKE ) TAKEDA CHEM IND LTD.
XX
XX
PI      Tanaka H, Kaieda I;
XX
XX
DR      WPI; 2002-435536/46.
XX
XX
DR      N-PSDB; AAL42859.
XX
XX
PT      Baculovirus IAP repeat domain or RING-finger domain-containing survivin-
PT      like polypeptides and encoded DNAs, applicable in diagnosis and screening
PT      compounds for treating various cancers and apoptosis abnormality.
XX
XX
PS      Disclosure; Page 122-123; 136pp; Japanese.
XX
XX
CC      The invention comprises the amino acid and coding sequences of survivin-
CC      like proteins. The survivin-like DNA and protein sequences are useful in
CC      diagnostics and screening compounds for treating various cancers and
CC      apoptosis abnormality, including gene therapy. The present amino acid
CC      sequence represents a survivin-like protein of the invention
XX
XX
SQ      Sequence 142 AA;

Query Match      100.0%; Score 778; DB 5; Length 142;
Best Local Similarity 100.0%; Pred. No. 7.9e-78;
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MGAPLPPAWQPFVKDHRISTFKNWPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60
        |||||||
Db      1 MGAPLPPAWQPFVKDHRISTFKNWPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60
        |||||||
QY      61 FKELEGWEPDDPIEHHKHSSGCAFLSVKKQFEEELTGEFLKLDREAKNKIAKETNNK 120
        |||||||
Db      61 FKELEGWEPDDPIEHHKHSSGCAFLSVKKQFEEELTGEFLKLDREAKNKIAKETNNK 120
        |||||||

```

QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 DB 121 KKEFEETAKKVRRAIEQLAAMD 142

## RESULT 6

ABP72162  
 ID ABP72162 standard; protein; 142 AA.

AC ABP72162;

DT 22-APR-2003 (first entry)

XX Human Survivin.

XX Human; Survivin; inhibitor of apoptosis; cell death; apoptosis; cancer;  
 KW cytostatic; cardiant; neuroprotective; gene therapy.

XX Homo sapiens.

PN WO2003004606-A2.

PD 16-JAN-2003.

XX 03-JUL-2002; 2002WO-US021002.

XX 03-JUL-2001; 2001US-00898158.

PA (UYCO ) UNIV COLUMBIA NEW YORK.

XX Troy CM, Shelanski ML;

XX WPI; 2003-210351/20.

DR N-PSDB; ABZ58106.

XX New nucleic acid encoding an inhibitor-of-apoptosis protein, useful for  
 PT treating cancer, neurodegenerative disorder or cardiomyopathy.

XX Disclosure; Fig 20A; 124pp; English.

XX The present sequence is the protein sequence of human Survivin protein.  
 CC The invention provides a nucleic acid, such as an antisense  
 CC oligonucleotide, which specifically hybridises to a nucleic acid encoding  
 CC an inhibitor of apoptosis protein. Survivin is an example of an inhibitor  
 CC of apoptosis or inducer of cell death protein of the invention. A claimed  
 CC method for inducing a cell's death comprises contacting the cell with the  
 CC nucleic acid under conditions permitting the nucleic acid to enter the  
 CC cell, especially the use of a vector, liposome, or a mechanical or  
 CC electrical means. The method is used to treat acute lymphocytic  
 CC leukaemia, acute myelogenous leukaemia, lung cancer, breast cancer,  
 CC ovarian cancer, prostate cancer, lymphoma, Hodgkin's disease, malignant  
 CC melanoma, neuroblastoma, renal cell carcinoma and squamous cell carcinoma  
 CC (all claimed). The invention also provides a second nucleic acid, which  
 CC specifically hybridises to a nucleic acid encoding a protein, other than  
 CC caspase-2, that induces cell death. A claimed method for inhibiting a  
 CC cell's death comprises contacting the cell with the nucleic acid under  
 CC conditions permitting the nucleic acid to enter the cell. The method is  
 CC used to treat a neurodegenerative disorder (especially a brain disorder  
 CC or central nervous system disorder), or a heart disorder (especially  
 CC cardiomyopathy) in a human (all claimed)

XX Sequence 142 AA;

Query Match 100.0%; Score 778; DB 6; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-78;  
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAPTLPPAWQPFLLKDHRISTFKNWPFLGCACTPERMAEAGFIHCTENEPDLAQCFCC 60

DB 1 MGAPTLPPAWQPFLLKDHRISTFKNWPFLGCACTPERMAEAGFIHCTENEPDLAQCFCC 60

QY 61 FKELEGWEPDDPPIEEHKKHSSGCAFLSVKKQFELTLGFLKLDRAKNKIATETNNK 120

DB 61 FKELEGWEPDDPPIEEHKKHSSGCAFLSVKKQFELTLGFLKLDRAKNKIATETNNK 120

QY 121 KKEFEETAKKVRRAIEQLAAMD 142

DB 121 KKEFEETAKKVRRAIEQLAAMD 142

## RESULT 7

ABR61582

ID ABR61582 standard; protein; 142 AA.

XX ABR61582;

XX 15-JAN-2004 (first entry)

XX Human survivin protein.

XX RasGAP; Aurora kinase; cytostatic; anti-tumoural; human; survivin.

XX Homo sapiens.

PN WO2003087395-A2.

PD 23-OCT-2003.

XX 15-APR-2003; 2003WO-IB002972.

XX 15-APR-2002; 2002US-0372483P.

PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

XX Garbay C, Gigoux V, Camonis J, L'hoste S, Samson J;

XX WPI; 2003-845340/78.

DR N-PSDB; ACF58071.

XX Identifying anti-tumoral compounds, comprises determining the capacity of  
 PT a compound to inhibit interaction between RasGAP and Drosophila  
 PT melanogaster (Dm) aurora kinase, or between RasGAP, Dm aurora kinase and  
 PT aurora binding protein.

XX Disclosure; Page 67; Opp; English.

XX The invention relates to identifying a biologically active compound with  
 CC anti-tumoural properties, where the compound is studied for its capacity  
 CC to inhibit the interaction between: (a) RasGAP and the Drosophila  
 CC melanogaster Aurora kinase or an orthologue of the kinase, or its  
 CC fragment able to interact with RasGAP; or (b) RasGAP, an Aurora-binding  
 CC protein and the D. melanogaster Aurora kinase or an orthologue of the  
 CC kinase, or its fragment. The method is useful for identifying anti-  
 CC tumoural agents. The present sequence represents a human survivin  
 CC protein, an aurora-binding protein

XX Sequence 142 AA;

Query Match 100.0%; Score 778; DB 7; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-78;  
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAPTLPPAWQPFLLKDHRISTFKNWPFLGCACTPERMAEAGFIHCTENEPDLAQCFCC 60

DB 1 MGAPTLPPAWQPFLLKDHRISTFKNWPFLGCACTPERMAEAGFIHCTENEPDLAQCFCC 60

QY 61 FKELEGWEPDDPPIEEHKKHSSGCAFLSVKKQFELTLGFLKLDRAKNKIATETNNK 120

DB 61 FKELEGWEPDDPPIEEHKKHSSGCAFLSVKKQFELTLGFLKLDRAKNKIATETNNK 120

QY 121 KKEFEETAKKVRRAIEQLAAMD 142

DB 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 8  
AAU29999  
ID AAU29999 standard; protein; 143 AA.  
XX  
AC AAU29999;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #490.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US008656.  
XX  
PR 18-APR-2000; 2000US-00552929.  
PR 26-JAN-2001; 2001US-00770160.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-611725/70.  
XX  
XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy.  
XX  
PS Claim 20; Page 219; 765pp; English.  
XX  
CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention  
XX  
SQ Sequence 143 AA;  
Query Match 100.0%; Score 778; DB 4; Length 143;  
Best Local Similarity 100.0%; Pred No. 7.9e-78;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGAPTLPPAQPFKDHRISTFKNWPFLGCACTPERMAEAGFIHCTPTNEPDLAQCFFC 60  
DB 1 MGAPTLPPAQPFKDHRISTFKNWPFLGCACTPERMAEAGFIHCTPTNEPDLAQCFFC 60  
QY 61 FKELEGWEPDDPIEEHKHSSGCAFLSVKKQFEELTLGFFLKDRERAKNKIATETNNK 120  
DB 61 FKELEGWEPDDPIEEHKHSSGCAFLSVKKQFEELTLGFFLKDRERAKNKIATETNNK 120  
QY 121 KGEFEETAKVRAIEQLAAMD 142  
DB 121 KGEFEETAKVRAIEQLAAMD 142

RESULT 9  
AAU74572  
ID AAU74572 standard; protein; 142 AA.  
XX  
AC AAU74572;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE Human mutant L54M survivin polypeptide.  
XX  
KW Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;  
KW dimerisation; ovary; breast; pancreas; central nervous system; blood;  
KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
KW cell proliferative disorder; acquired immunodeficiency syndrome;  
KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
KW myelodysplastic syndrome; protein coordinate data; mutant; mutein.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 54  
FT /note= "Wild-type Leu substituted by Met"  
XX  
PN WO200202622-A2.  
XX  
PD 10-JAN-2002.  
XX  
PF 29-JUN-2001; 2001WO-US020872.  
XX  
PR 29-JUN-2000; 2000US-00608352.  
XX  
PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
XX  
PI Noel JB, Verdacia M, Hunter T, Huang H;  
XX  
DR WPI; 2002-171638/22.  
XX  
PT Novel isolated crystalline survivin polypeptide, useful to screen and  
PT design compounds that bind to or interact with inhibitor of apoptosis  
PT protein and protein family members, and for design of novel therapeutics.  
XX  
PS Claim 7; Page; 102pp; English.  
XX  
CC The invention relates to an isolated crystalline human survivin  
CC polypeptide, an inhibitor of apoptosis protein (IAP). The polypeptide is  
CC useful for determining the ability of a survivin binding agent to  
CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc  
CC chelation activity and for identifying an agent such as a peptide or  
CC peptidomimetic which inhibits dimerisation of survivin. The ability of an  
CC agent to modulate dimerisation can be determined through detection of a  
CC change in apoptosis in a target cell expressing survivin. Binding agents  
CC are useful for increasing apoptosis in a cell derived from a tissue  
CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,  
CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous  
CC system, prostate, colon, rectum, cervix or endometrium, with a cell  
CC proliferative disorder such as cancer. The molecules can also be used in  
CC treatment or prevention of apoptosis which occurs as a part of AIDS,  
CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease  
CC and myelodysplastic syndromes. This sequence represents a human mutant  
CC survivin polypeptide. Note: This sequence is not featured in the printed  
CC specification but was derived from the wild type protein shown in  
XX  
SQ Sequence 142 AA;

Query Match 99.7%; Score 776; DB 5; Length 142;  
Best Local Similarity 99.3%; Pred. No. 1.3e-77;  
Matches 141; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGAPTLPPAQPFKDHRISTFKNWPFLGCACTPERMAEAGFIHCTPTNEPDLAQCFFC 60

Db 1 MGAPTLPPAWQFLKDHRISTFKWPFLEGGCACTPERMAEAGFIHCPTENPDLAQCFFC 60  
 QY 61 FKELEGWEPDDPIEHHKHSAGCAFLSVKKQFBEITLGEFLKDRBRANKIAKETNNK 120  
 Db 61 FKELEGWEPDDPIEHHKHSAGCAFLSVKKQFBEITLGEFLKDRBRANKIAKETNNK 120  
 QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 Db 121 KKEFEETAKKVRRAIEQLAAMD 142  
 RESULT 10  
 AAU74574  
 ID AAU74574 standard; protein; 142 AA.  
 AC AAU74574;  
 XX  
 XX 08-MAY-2002 (first entry)  
 DT Human mutant T34E survivin polypeptide.  
 DE  
 DE Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
 KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;  
 KW dimerisation; ovary; breast; pancreas; central nervous system; blood;  
 KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
 KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
 KW cell proliferative disorder; acquired immunodeficiency syndrome;  
 KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
 KW myelodysplastic syndrome; protein coordinate data; mutant; mutein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 34 /note= "Wild-type Thr substituted by Glu"  
 FT  
 FN WO200202622-A2.  
 PD 10-JAN-2002.  
 XX  
 XX 29-JUN-2001; 2001WO-US020872.  
 XX  
 XX 29-JUN-2000; 2000US-00608352.  
 XX  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX  
 XX Noel JP, Verdacia M, Hunter T, Huang H;  
 XX WPI; 2002-171638/22.  
 XX  
 XX Novel isolated crystalline survivin polypeptide, useful to screen and  
 PT design compounds that bind to or interact with inhibitor of apoptosis  
 PT protein and protein family members, and for design of novel therapeutics.  
 XX  
 XX Claim 8; Page; 102pp; English.  
 XX  
 XX The invention relates to an isolated crystalline human survivin  
 CC polypeptide, an inhibitor of apoptosis protein (IAP). The polypeptide is  
 CC useful for determining the ability of a survivin binding agent to  
 CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc  
 CC chelation activity and for identifying an agent such as a peptide or  
 CC peptidomimetic which inhibits dimerisation of survivin. The ability of an  
 CC agent to modulate dimerisation can be determined through detection of a  
 CC change in apoptosis in a target cell expressing survivin. Binding agents  
 CC are useful for increasing apoptosis in a cell derived from a tissue  
 CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,  
 CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous  
 CC system, prostate, colon, rectum, cervix or endometrium, with a cell  
 CC proliferative disorder such as cancer. The molecules can also be used in  
 CC treatment or prevention of apoptosis which occurs as a part of AIDS,  
 CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease

CC and myelodysplastic syndromes. This sequence represents a human mutant  
 CC survivin polypeptide. Note: This sequence is not featured in the printed  
 CC specification but was derived from the wild type protein shown in  
 CC AAU74571  
 XX  
 SQ Sequence 142 AA;  
 Query Match 99.2%; Score 772; DB 5; Length 142;  
 Best Local Similarity 99.3%; Pred. No. 3.6e-77;  
 Matches 141; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGAPTLPPAWQFLKDHRISTFKWPFLEGGCACTPERMAEAGFIHCPTENPDLAQCFFC 60  
 Db 1 MGAPTLPPAWQFLKDHRISTFKWPFLEGGCACTPERMAEAGFIHCPTENPDLAQCFFC 60  
 QY 61 FKELEGWEPDDPIEHHKHSAGCAFLSVKKQFBEITLGEFLKDRBRANKIAKETNNK 120  
 Db 61 FKELEGWEPDDPIEHHKHSAGCAFLSVKKQFBEITLGEFLKDRBRANKIAKETNNK 120  
 QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 Db 121 KKEFEETAKKVRRAIEQLAAMD 142  
 RESULT 11  
 AAU74577  
 ID AAU74577 standard; protein; 142 AA.  
 XX  
 XX AAU74577;  
 AC  
 XX 08-MAY-2002 (first entry)  
 DT Human mutant T97E survivin polypeptide.  
 DE  
 DE Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
 KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;  
 KW dimerisation; ovary; breast; pancreas; central nervous system; blood;  
 KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
 KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
 KW cell proliferative disorder; acquired immunodeficiency syndrome;  
 KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
 KW myelodysplastic syndrome; protein coordinate data; mutant; mutein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 97 /note= "Wild-type Thr substituted by Glu"  
 FT  
 FN WO200202622-A2.  
 PD 10-JAN-2002.  
 XX  
 XX 29-JUN-2001; 2001WO-US020872.  
 XX  
 XX 29-JUN-2000; 2000US-00608352.  
 XX  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX  
 XX Noel JP, Verdacia M, Hunter T, Huang H;  
 XX WPI; 2002-171638/22.  
 XX  
 XX Novel isolated crystalline survivin polypeptide, useful to screen and  
 PT design compounds that bind to or interact with inhibitor of apoptosis  
 PT protein and protein family members, and for design of novel therapeutics.  
 XX  
 XX Claim 8; Page; 102pp; English.  
 XX  
 XX The invention relates to an isolated crystalline human survivin  
 CC polypeptide, an inhibitor of apoptosis protein (IAP). The polypeptide is  
 CC useful for determining the ability of a survivin binding agent to

CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc  
 CC chelation activity and for identifying an agent such as a peptide or  
 CC peptidomimetic which inhibits dimerisation of survivin. The ability of an  
 CC agent to modulate dimerisation can be determined through detection of a  
 CC change in apoptosis in a target cell expressing survivin. Binding agents  
 CC are useful for increasing apoptosis in a cell derived from a tissue  
 CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,  
 CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous  
 CC system, prostate, colon, rectum, cervix or endometrium, with a cell  
 CC proliferative disorder such as cancer. The molecules can also be used in  
 CC treatment or prevention of apoptosis which occurs as a part of AIDS,  
 CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease  
 CC and myelodysplastic syndromes. This sequence represents a human mutant  
 CC survivin polypeptide. Note: This sequence is not featured in the printed  
 CC specification but was derived from the wild type protein shown in  
 CC AAU74571  
 XX  
 XX Sequence 142 AA;

Query Match 99.2%; Score 772; DB 5; Length 142;  
 Best Local Similarity 99.3%; Pred. No. 3.6e-77;  
 Matches 141; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGAPTLPPAWQPFLLKDHRISTFKNWPFLGCGACTPERMAEAGFIHCTENEPDLAQCFPC 60  
 Db 1 MGAPTLPPAWQPFLLKDHRISTFKNWPFLGCGACTPERMAEAGFIHCTENEPDLAQCFPC 60  
 QY 61 FKELEGWEPDDDDPIEHHKSSGCAFLSVKKQFEELTGLGFLKLDREAKNKIAKETNNK 120  
 Db 61 FKELEGWEPDDDDPIEHHKSSGCAFLSVKKQFEELTGLGFLKLDREAKNKIAKETNNK 120  
 QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 Db 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 12  
 AAU74575  
 ID AAU74575 standard; protein; 142 AA.

AC AAU74575;  
 XX  
 DT 08-MAY-2002 (first entry)  
 DE Human mutant H80A survivin polypeptide.  
 DE  
 KW Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
 KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;  
 KW dimerisation; ovary; breast; pancreas; central nervous system; blood;  
 KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
 KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
 KW cell proliferative disorder; acquired immunodeficiency syndrome;  
 KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
 KW myelodysplastic syndrome; protein coordinate data; mutant; muten.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.

XX Key Location/Qualifiers  
 XX Misc-difference 80  
 FT /note= "Wild-type His substituted by Ala"

XX WO200202622-A2.  
 XX 10-JAN-2002.  
 XX 29-JUN-2001; 2001WO-US020872.  
 XX 29-JUN-2000; 2000US-00608352.  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.

XX Noel JP, Verdacia M, Hunter T, Huang H;  
 PI

XX WPI; 2002-171638/22.

XX Novel isolated crystalline survivin polypeptide, useful to screen and  
 PT design compounds that bind to or interact with inhibitor of apoptosis  
 PT protein and protein family members, and for design of novel therapeutics.

XX Claim 8; Page; 102pp; English.

XX The invention relates to an isolated crystalline human survivin  
 CC polypeptide, an inhibitor of apoptosis protein (IAP). The polypeptide is  
 CC useful for determining the ability of a survivin binding agent to  
 CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc  
 CC chelation activity and for identifying an agent such as a peptide or  
 CC peptidomimetic which inhibits dimerisation of survivin. The ability of an  
 CC agent to modulate dimerisation can be determined through detection of a  
 CC change in apoptosis in a target cell expressing survivin. Binding agents  
 CC are useful for increasing apoptosis in a cell derived from a tissue  
 CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,  
 CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous  
 CC system, prostate, colon, rectum, cervix or endometrium, with a cell  
 CC proliferative disorder such as cancer. The molecules can also be used in  
 CC treatment or prevention of apoptosis which occurs as a part of AIDS,  
 CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease  
 CC and myelodysplastic syndromes. This sequence represents a human mutant  
 CC survivin polypeptide. Note: This sequence is not featured in the printed  
 CC specification but was derived from the wild type protein shown in  
 CC AAU74571

XX Sequence 142 AA;

Query Match 98.7%; Score 768; DB 5; Length 142;  
 Best Local Similarity 99.3%; Pred. No. 1e-76;  
 Matches 141; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGAPTLPPAWQPFLLKDHRISTFKNWPFLGCGACTPERMAEAGFIHCTENEPDLAQCFPC 60  
 Db 1 MGAPTLPPAWQPFLLKDHRISTFKNWPFLGCGACTPERMAEAGFIHCTENEPDLAQCFPC 60  
 QY 61 FKELEGWEPDDDDPIEHHKSSGCAFLSVKKQFEELTGLGFLKLDREAKNKIAKETNNK 120  
 Db 61 FKELEGWEPDDDDPIEHHKSSGCAFLSVKKQFEELTGLGFLKLDREAKNKIAKETNNK 120  
 QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 Db 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 13  
 AAU74573  
 ID AAU74573 standard; protein; 142 AA.

XX AAU74573;  
 XX AC  
 XX 08-MAY-2002 (first entry)  
 DT  
 DE Human mutant W10A survivin polypeptide.

XX Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
 KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;  
 KW dimerisation; ovary; breast; pancreas; central nervous system; blood;  
 KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
 KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
 KW cell proliferative disorder; acquired immunodeficiency syndrome;  
 KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
 KW myelodysplastic syndrome; protein coordinate data; mutant; muten.

XX Homo sapiens.  
 OS Synthetic.

XX Key Location/Qualifiers  
 XX Misc-difference 10  
 FT /note= "Wild-type Trp substituted by Ala"

XX WO200202622-A2.  
 XX 10-JAN-2002.  
 XX 29-JUN-2001; 2001WO-US020872.  
 XX 29-JUN-2000; 2000US-00608352.  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX Noel JP, Verdacia M, Hunter T, Huang H;  
 XX WPI; 2002-171638/22.  
 XX Novel isolated crystalline survivin polypeptide, useful to screen and  
 PT design compounds that bind to or interact with inhibitor of apoptosis  
 PT protein and protein family members, and for design of novel therapeutics.  
 XX  
 XX Claim 8; Page; 102pp; English.  
 XX  
 CC The invention relates to an isolated crystalline human survivin  
 CC polypeptide, an inhibitor or apoptosis protein (IAP). The polypeptide is  
 CC useful for determining the ability of a survivin binding agent to  
 CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc  
 CC chelation activity and for identifying an agent such as a peptide or  
 CC peptidomimetic which inhibits dimerisation of survivin. The ability of an  
 CC agent to modulate dimerisation can be determined through detection of a  
 CC change in apoptosis in a target cell expressing survivin. Binding agents  
 CC are useful for increasing apoptosis in a cell derived from a tissue  
 CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,  
 CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous  
 CC system, prostate, colon, rectum, cervix or endometrium, with a cell  
 CC proliferative disorder such as cancer. The molecules can also be used in  
 CC treatment or prevention of apoptosis which occurs as a part of AIDS,  
 CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease  
 CC and myelodysplastic syndromes. This sequence represents a human mutant  
 CC survivin polypeptide. Note: This sequence is not featured in the printed  
 CC specification but was derived from the wild type protein shown in  
 CC AAU74571  
 XX  
 XX Sequence 142 AA;  
 SQ  
 Query Match 98.2%; Score 764; DB 5; Length 142;  
 Best Local Similarity 99.3%; Pred. No. 2.8e-76;  
 Matches 141; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGAPTLPPAWQPFLLKDHRISTFKWPFLEGGCACTPERMAEAGFTHCPTNEPDLAQCFFC 60  
 DB 1 MGAPTLPPAAQPFLLKDHRISTFKWPFLEGGCACTPERMAEAGFTHCPTNEPDLAQCFFC 60  
 QY 61 FKELEGWEPDDPIEEHKHSGGCAFLSVKKQFELTLGFLKLDRAKNIKETNNK 120  
 DB 61 FKELEGWEPDDPIEEHKHSGGCAFLSVKKQFELTLGFLKLDRAKNIKETNNK 120  
 QY 121 KKEFEETAKVRRATEQLAAMD 142  
 DB 121 KKEFEETAKVRRATEQLAAMD 142  
 QY 121 KKEFEETAKVRRATEQLAAMD 142  
 DB 121 KKEFEETAKVRRATEQLAAMD 142  
 RESULT 14  
 ABUS6477  
 ID ABUS6477 standard; protein; 141 AA.  
 XX  
 XX AC ABUS6477;  
 XX  
 XX 02-APR-2003 (first entry)  
 DE Lung cancer-associated polypeptide #70.  
 XX Lung cancer-associated polypeptide; cytostatic; emphysema;  
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
 XX Unidentified.  
 XX WO2002026443-A2.  
 XX 31-OCT-2002.  
 XX 18-APR-2002; 2002WO-US012476.  
 XX 18-APR-2001; 2001US-0284770P.  
 PR 10-MAY-2001; 2001US-0290492P.  
 PR 09-NOV-2001; 2001US-0339245P.  
 PR 13-NOV-2001; 2001US-0350666P.  
 PR 29-NOV-2001; 2001US-0334370P.  
 PR 12-APR-2002; 2002US-0372246P.  
 XX (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 XX Aziz N, Murray R;  
 XX WPI; 2003-093161/08.  
 DR N-PSDB; ABX76199.  
 DR  
 PT Detecting a lung cancer-associated transcript in a cell from a patient  
 PT for treating lung cancer, by contacting a biological sample from the  
 PT patient with a polynucleotide that exhibits increased or decreased  
 PT expression in lung cancer.  
 XX  
 PS Claim 27; Page 242; 453pp; English.  
 XX  
 CC The invention relates to a method for detecting a lung cancer-associated  
 CC transcript in a cell from a patient, comprising contacting a biological  
 CC sample from the patient with a polynucleotide that selectively hybridises  
 CC to a sequence that is at least 80 % identical to a gene that exhibits  
 CC increased or decreased expression in lung cancer samples. Lung cancer-  
 CC associated polynucleotides and polypeptides are used for identifying a  
 CC compound that modulates a lung cancer-associated polypeptide, for  
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung  
 CC cancer in a patient and for treating a mammal having lung cancer by  
 CC administering a modulatory compound identified. The methods are useful  
 CC for treating lung cancer, such as small cell lung cancer, non-small cell  
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,  
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,  
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and  
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful  
 CC for diagnostic purposes and as targets for screening for therapeutic  
 CC compounds that modulate lung cancer, such as antibodies. Sequences  
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the  
 CC invention  
 XX  
 XX Sequence 141 AA;  
 SQ  
 Query Match 98.0%; Score 762.5; DB 6; Length 141;  
 Best Local Similarity 99.3%; Pred. No. 4e-76;  
 Matches 141; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 MGAPTLPPAWQPFLLKDHRISTFKWPFLEGGCACTPERMAEAGFTHCPTNEPDLAQCFFC 60  
 DB 1 MGAPTLPPAWQPFLLKDHRISTFKWPFLEGGCACTPERMAEAGFTHCPTNEPDLAQCFFC 60  
 QY 61 FKELEGWEPDDPIEEHKHSGGCAFLSVKKQFELTLGFLKLDRAKNIKETNNK 120  
 DB 61 FKELEGWEPDDPIEEHKHSGGCAFLSVKKQFELTLGFLKLDRAKNIKETNNK 120  
 QY 121 KKEFEETAKVRRATEQLAAMD 142  
 DB 121 KKEFEETAKVRRATEQLAAMD 141  
 RESULT 15  
 AAU74576



ID AAU74576 standard; protein; 142 AA.  
AC AAU74576;  
XX  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE Human mutant H80A/E76A survivin polypeptide.  
XX  
XX Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;  
KW dimerisation; ovary; breast; pancreas; central nervous system; blood;  
KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
KW cell proliferative disorder; acquired immunodeficiency syndrome;  
KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
KW myelodysplastic syndrome; protein coordinate data; mutant; mutuin.  
XX  
OS Homo sapiens.  
OS Synthetic.  
OS  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 76 /note= "Wild-type Glu substituted by Ala"  
FT FT /note= "Wild-type Glu substituted by Ala"  
FT Misc-difference 80 /note= "Wild-type His substituted by Ala"  
XX  
XX WO200202622-A2.  
XX  
XX 10-JAN-2002.  
XX  
XX 29-JUN-2001; 2001WO-US020872.  
XX  
XX 29-JUN-2000; 2000US-00608352.  
XX  
XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
XX  
XX Noel JP, Verdacia M, Hunter T, Huang H;  
PI WPI; 2002-171638/22.  
XX  
XX Novel isolated crystalline survivin polypeptide, useful to screen and  
FT design compounds that bind to or interact with inhibitor of apoptosis  
FT protein and protein family members, and for design of novel therapeutics.  
XX  
XX Claim 8; Page; 102pp; English.  
XX  
XX The invention relates to an isolated crystalline human survivin  
CC polypeptide, an inhibitor or apoptosis protein (IAP). The polypeptide is  
CC useful for determining the ability of a survivin binding agent to  
CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc  
CC chelation activity and for identifying an agent such as a peptide or  
CC peptidomimetic which inhibits dimerisation of survivin. The ability of an  
CC agent to modulate dimerisation can be determined through detection of a  
CC change in apoptosis in a target cell expressing survivin. Binding agents  
CC are useful for increasing apoptosis in a cell derived from a tissue  
CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,  
CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous  
CC system, prostate, colon, rectum, cervix or endometrium, with a cell  
CC proliferative disorder such as cancer. The molecules can also be used in  
CC treatment or prevention of apoptosis which occurs as a part of AIDS,  
CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease  
CC and myelodysplastic syndromes. This sequence represents a human mutant  
CC survivin polypeptide. Note: This sequence is not featured in the printed  
CC specification but was derived from the wild type protein shown in  
XX AAU74571  
XX  
XX Sequence 142 AA;  
XX  
XX Query Match 97.9%; Score 762; DB 5; Length 142;  
XX Best Local Similarity 98.6%; Pred. No. 4.6e-76;  
XX Matches 140; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
XX  
XX 1 MGAPILPFAWQFFLKDHRISTFKWPFLEGGCACTPERVAEAGFIHCPTENEPDLAQCFPC 60

Db 1 MGAPILPFAWQFFLKDHRISTFKWPFLEGGCACTPERVAEAGFIHCPTENEPDLAQCFPC 60  
QY 61 FKELEGWEPDDDDPIEEHKKHSSGCAFLSVKKQFBEFLTGLKLDREKAKNIAKETNNK 120  
Db 61 FKELEGWEPDDDDPIEAHKKASSGCAFLSVKKQFBEFLTGLKLDREKAKNIAKETNNK 120  
QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
Db 121 KKEFEETAKKVRRAIEQLAAMD 142

Search completed: August 11, 2004, 14:15:54  
Job time : 54 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw\_model

Run on: August 11, 2004, 14:17:11 ; Search time 46 Seconds  
(without alignments)  
969.079 Million cell updates/sec

Title: US-09-690-825-34  
Perfect score: 778  
Sequence: 1 MGAPLPPAWQPLKDHRS.....EFBETAKKVRRAIEQLAAMD 142

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	778	100.0	142	14	US-10-138-618-34
2	778	100.0	142	16	US-10-665-975-2
3	774	99.5	142	16	US-10-408-765A-2210
4	562	72.2	98	9	US-09-965-967-23
5	155	19.9	1403	8	US-08-913-322-22
6	155	19.9	1403	8	US-08-913-322-24
7	155	19.9	1403	14	US-10-285-408-1
8	151.5	19.5	172	14	US-10-041-859-8
9	148.5	19.1	172	14	US-10-041-859-10
10	145.5	18.7	172	14	US-10-041-859-11
11	144.5	18.6	172	14	US-10-041-859-12
12	143	18.4	346	14	US-10-041-859-2
13	139.5	17.9	127	12	US-10-424-539-246508
14	138.5	17.8	172	14	US-10-041-859-9
15	138	17.7	355	14	US-10-203-708-44

Sequence 2, Appli  
Sequence 14, Appli  
Sequence 12, Appli  
Sequence 9, Appli  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 10, Appli  
Sequence 29, Appli  
Sequence 17, Appli  
Sequence 28, Appli  
Sequence 16, Appli  
Sequence 18, Appli  
Sequence 8, Appli  
Sequence 3, Appli  
Sequence 63, Appli  
Sequence 4, Appli  
Sequence 338, App  
Sequence 200, App  
Sequence 2, Appli  
Sequence 223, App  
Sequence 2, Appli  
Sequence 20, Appli  
Sequence 10, Appli  
Sequence 10, Appli  
Sequence 225, App  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 23, Appli  
Sequence 34, Appli

464 12 US-10-343-115-2  
68 14 US-10-041-859-14  
275 9 US-09-201-936-12  
275 14 US-10-323-643-9  
280 14 US-10-248-586-3  
298 14 US-10-235-026-2  
268 14 US-10-323-643-10  
68 9 US-09-201-936-29  
68 14 US-10-041-859-17  
68 9 US-09-201-936-28  
68 14 US-10-041-859-16  
68 14 US-10-041-859-18  
618 9 US-09-974-592-8  
618 9 US-09-201-936-8  
618 12 US-10-361-270-3  
618 12 US-10-260-708-63  
618 12 US-10-366-307-4  
618 14 US-10-153-668-338  
618 14 US-10-207-658-200  
618 14 US-10-232-286-2  
618 16 US-10-636-065-223  
236 12 US-10-024-433-2  
68 9 US-09-201-936-20  
496 9 US-09-974-592-10  
496 9 US-09-201-936-10  
496 16 US-10-636-085-225  
612 9 US-09-974-592-14  
612 14 US-10-232-286-14  
67 9 US-09-201-936-23  
148 10 US-09-764-861-34

## ALIGNMENTS

RESULT 1  
US-10-138-618-34  
; Sequence 34, Application US/10138618  
; Publication No. US20030100525A1  
; GENERAL INFORMATION:  
; APPLICANT: Altieri, Dario C.  
; TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS CELLULAR APOPTOSIS, AND ITS MODULATION

NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
STREET: 1800 M Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036-5869

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION NUMBER:  
APPLICATION NUMBER: US/10138,618  
FILING DATE: 06-May-2002

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/975,080  
FILING DATE: 20-NOV-1997  
APPLICATION NUMBER: US 60/031,435  
FILING DATE: 20-NOV-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Reid G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 044574-5022-01-WO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-467-7000  
TELEFAX: 202-467-7176  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:

```
;
; LENGTH: 142 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-138-618-34

Query Match          100.0%; Score 778; DB 14; Length 142;
Best Local Similarity 100.0%; Pred. No. 1.1e-72;
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAPTLPPAWQPFLLKDHRISTFFKNWPFLEGCACCTPRMAEAGFIHCPTENEPDLAQCFPC 60
DB 1 MGAPTLPPAWQPFLLKDHRISTFFKNWPFLEGCACCTPRMAEAGFIHCPTENEPDLAQCFPC 60

QY 61 FKELEGWEPDDDDPIEEHKHSSGCAFLSVKKQFEEELTLGFLKLDRAKKNIAKETNNK 120
DB 61 FKELEGWEPDDDDPIEEHKHSSGCAFLSVKKQFEEELTLGFLKLDRAKKNIAKETNNK 120

QY 121 KKEFEETAKKVRRAIEQLAAMD 142
DB 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 2
US-10-665-975-2
; Sequence 2, Application US/10665975
; Publication No. US20040138119A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Tamm, Ingo
; TITLE OF INVENTION: USE OF HEPATITIS VIRUS B X-INTERACTING
; FILE REFERENCE: BURNHAM.C05A
; CURRENT APPLICATION NUMBER: US/10/665,975
; CURRENT FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/412,109
; PRIOR FILING DATE: 2002-09-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-975-2

Query Match          100.0%; Score 778; DB 16; Length 142;
Best Local Similarity 100.0%; Pred. No. 1.1e-72;
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAPTLPPAWQPFLLKDHRISTFFKNWPFLEGCACCTPRMAEAGFIHCPTENEPDLAQCFPC 60
DB 1 MGAPTLPPAWQPFLLKDHRISTFFKNWPFLEGCACCTPRMAEAGFIHCPTENEPDLAQCFPC 60

QY 61 FKELEGWEPDDDDPIEEHKHSSGCAFLSVKKQFEEELTLGFLKLDRAKKNIAKETNNK 120
DB 61 FKELEGWEPDDDDPIEEHKHSSGCAFLSVKKQFEEELTLGFLKLDRAKKNIAKETNNK 120

QY 121 KKEFEETAKKVRRAIEQLAAMD 142
DB 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 3
US-10-408-765A-2210
; Sequence 2210, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Rahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
```

```
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2210
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2210

Query Match          99.5%; Score 774; DB 16; Length 142;
Best Local Similarity 99.3%; Pred. No. 2.8e-72;
Matches 141; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAPTLPPAWQPFLLKDHRISTFFKNWPFLEGCACCTPRMAEAGFIHCPTENEPDLAQCFPC 60
DB 1 MGAPTLPPAWQPFLLKDHRISTFFKNWPFLEGCACCTPRMAEAGFIHCPTENEPDLAQCFPC 60

QY 61 FKELEGWEPDDDDPIEEHKHSSGCAFLSVKKQFEEELTLGFLKLDRAKKNIAKETNNK 120
DB 61 FKELEGWEPDDDDPIEEHKHSSGCAFLSVKKQFEEELTLGFLKLDRAKKNIAKETNNK 120

QY 121 KKEFEETAKKVRRAIEQLAAMD 142
DB 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 4
US-09-965-967-23
; Sequence 23, Application US/09965967
; Patent No. US20020177557A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Yigong
; TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
; FILE REFERENCE: PU-0031 (01-1739-1)
; CURRENT APPLICATION NUMBER: US/09/965,967
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,574
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/256,830
; PRIOR FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-967-23

Query Match          72.2%; Score 562; DB 9; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.8e-50;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAPTLPPAWQPFLLKDHRISTFFKNWPFLEGCACCTPRMAEAGFIHCPTENEPDLAQCFPC 60
DB 1 MGAPTLPPAWQPFLLKDHRISTFFKNWPFLEGCACCTPRMAEAGFIHCPTENEPDLAQCFPC 60

QY 61 FKELEGWEPDDDDPIEEHKHSSGCAFLSVKKQFEEELTL 98
DB 61 FKELEGWEPDDDDPIEEHKHSSGCAFLSVKKQFEEELTL 98

RESULT 5
US-08-913-322-22
; Sequence 22, Application US/08913322
; Publication No. US20020137028A1
; GENERAL INFORMATION:
```





Db 240 ARWDRCAVYLVKGRDVIQ-----KVKSEATAISASEEQAATNDSTKNVAQEGEK 291

RESULT 13

US-10-424-599-246508

; Sequence 246508, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalick David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 246508

; LENGTH: 127

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_64628C.1.pep

US-10-424-599-246508

Query Match 17.9%; Score 139.5; DB 12; Length 127;

Best Local Similarity 27.0%; Pred. No. 1.7e-06;

Matches 31; Conservative 21; Mismatches 46; Indels 17; Gaps 3;

Qy 18 RISTFK---NWPLEGCACTPERMAEAGFIHCPTENEPDLAQCFKLEGWEP-DDD 72

Db 10 RIKSKGRKKNPKYKRDTPVLDKLAAGFIHKPLSNHTDNVECFLCETAVNGWSDHND 69

Qy 73 PIEBKHSSGCAFLSVKKQFEELTGLFLKDRKAKKIATNKKKEFEET 127

Db 70 PFRRHIOHAPACAWML-----ICVARGRPVSLDSSGENVQNNQFDST 112

RESULT 14

US-10-041-859-9

; Sequence 9, Application US/10041859

; Publication No. US20030049796A1

; GENERAL INFORMATION:

; APPLICANT: HUANG, QIHONG

; APPLICANT: REED, JOHN C.

; APPLICANT: DEVERAUX, QUINN L.

; APPLICANT: MAEDA, SUSUMU

; TITLE OF INVENTION: INHIBITOR OF APOPTOSIS PROTEINS AND NUCLEIC ACIDS AND

; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM

; FILE REFERENCE: 087102/027 2537

; CURRENT APPLICATION NUMBER: US/10/041,859

; CURRENT FILING DATE: 2002-01-07

; PRIOR APPLICATION NUMBER: 60/260,478

; PRIOR FILING DATE: 2001-01-08

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 9

; LENGTH: 172

; TYPE: PRT

; ORGANISM: Spodoptera frugiperda

US-10-041-859-9

Query Match 17.8%; Score 138.5; DB 14; Length 172;

Best Local Similarity 34.5%; Pred. No. 3.2e-06;

Matches 30; Conservative 16; Mismatches 30; Indels 11; Gaps 3;

Qy 12 PFLKDHRISTPKNWPFLGGAC---TPERMAEAGFIHCPTENEPDLAQCFKLEGW 68

Db 65 PFVEAARLSPKDWF-----RCWQKPELAAGFFY---TGQGDNTKCFYCDGGLKWE 116

Qy 69 PDDPIEBHKKHSSGCAFLSVKKQFEE 95

Db 117 NHDVPWEQHARWFDRCAYVLCIKCYAE 143

RESULT 15

US-10-203-708-44

; Sequence 44, Application US/10203708

; Publication No. US20030149238A1

; GENERAL INFORMATION:

; APPLICANT: SMITHKLINE BEECHAM CORPORATION

; APPLICANT: SMITHKLINE BEECHAM P.L.C.

; TITLE OF INVENTION: NOVEL COMPOUNDS

; FILE REFERENCE: GP50013

; CURRENT APPLICATION NUMBER: US/10/203,708

; CURRENT FILING DATE: 2002-08-13

; PRIOR APPLICATION NUMBER: PCT/US01/04703

; PRIOR FILING DATE: 2001-02-14

; PRIOR APPLICATION NUMBER: 60/182,172

; PRIOR FILING DATE: 2000-02-14

; PRIOR APPLICATION NUMBER: 60/186,084

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 44

; LENGTH: 355

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-203-708-44

Query Match 17.7%; Score 138; DB 14; Length 355;

Best Local Similarity 32.7%; Pred. No. 8.7e-06;

Matches 33; Conservative 14; Mismatches 38; Indels 16; Gaps 4;

Qy 18 RISTFKNWPFLGCACTPERMAEAGFIHCPTENEPDLAQCFKLEGWEPDDDDPIEEH 77

Db 90 RLASFVDWELT--AEVFPPELLAAGFFH---TGHQDKVRCFCYCGLGSLQSKRGDDPWTEH 144

Qy 78 KHSSGCAFL-----SVKKQFEELTGLFLKDRER 108

Db 145 AKWFPSCQFLLRKSGRDFVHSVQETHSQL-LGSWYSATSPR 184

Search completed: August 11, 2004, 14:22:45

Job time : 47 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 14:14:55 ; Search time 19 Seconds  
(without alignments)  
385.836 Million cell updates/sec

Title: US-09-690-825-34  
Perfect score: 778  
Sequence: 1 MGAPTLPPAQWQPFKDHRISTFKNWFFLEGCACTPERMAEAGFIHCPTNEPDLAQCFFC 60

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/ptodata/2/iaa/PCRIUS COMB.pcp.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	100.0	142	3	US-08-975-080-34
2	778	100.0	142	4	US-09-283-144-3
3	246.5	31.7	153	4	US-09-738-946-4
4	219	28.1	137	4	US-09-283-144-15
5	155	19.9	1151	3	US-08-836-134-23
6	155	19.9	1151	4	US-09-493-784-23
7	155	19.9	1232	3	US-08-836-134-2
8	155	19.9	1232	4	US-09-493-784-2
9	155	19.9	1295	4	US-09-705-872-3
10	155	19.9	1403	4	US-09-705-872-1
11	140	18.0	75	4	US-08-557-759-22
12	136.5	17.5	275	2	US-08-511-485-12
13	136.5	17.5	275	3	US-08-836-134-21
14	136.5	17.5	275	4	US-09-493-784-21
15	136.5	17.5	275	4	US-09-201-936-12
16	136.5	17.5	275	4	US-09-011-356-12
17	136.5	17.5	277	4	US-09-127-928-2
18	136.5	17.5	377	4	US-09-502-528-3
19	133.5	17.2	268	3	US-08-836-134-22
20	133.5	17.2	268	4	US-09-493-784-22
21	132.5	17.0	68	2	US-08-511-485-29
22	132.5	17.0	68	4	US-09-201-936-29
23	132.5	17.0	68	4	US-09-011-356-29
24	131.5	16.9	68	2	US-08-511-485-28
25	131.5	16.9	68	4	US-09-201-936-28
26	131.5	16.9	68	4	US-09-011-356-28
27	131	16.8	618	2	US-08-511-485-8

28	131	16.8	618	3	US-09-212-971-8	Sequence 8, Appl
29	131	16.8	618	3	US-08-800-929A-8	Sequence 8, Appl
30	131	16.8	618	3	US-08-569-749-2	Sequence 2, Appl
31	131	16.8	618	4	US-09-617-053A-8	Sequence 8, Appl
32	131	16.8	618	4	US-09-069-023-29	Sequence 29, Appl
33	131	16.8	618	4	US-09-201-936-8	Sequence 8, Appl
34	131	16.8	618	4	US-09-011-356-8	Sequence 8, Appl
35	131	16.8	618	4	US-09-572-717-223	Sequence 223, App
36	131	16.8	618	5	PCT-US96-12860-2	Sequence 2, Appl
37	130	16.7	236	3	US-09-121-979-4	Sequence 4, Appl
38	130	16.7	236	3	US-09-332-319-4	Sequence 2, Appl
39	130	16.7	236	4	US-09-239-867-2	Sequence 20, Appl
40	129.5	16.6	68	2	US-08-511-485-20	Sequence 20, Appl
41	129.5	16.6	68	4	US-09-201-936-20	Sequence 20, Appl
42	129.5	16.6	68	4	US-09-011-356-20	Sequence 20, Appl
43	129.5	16.6	496	2	US-08-511-485-10	Sequence 10, Appl
44	129.5	16.6	496	3	US-09-212-971-10	Sequence 10, Appl
45	129.5	16.6	496	3	US-08-800-929A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-08-975-080-34  
; Sequence 34, Application US/08975080  
; Patent No. 6245523  
; GENERAL INFORMATION:  
; APELLICANT: Altieri, Dario C.  
; TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS  
; CELLULAR APOPTOSIS, AND ITS MODULATION  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
; STREET: 1800 M Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036-5869  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/975,080  
; FILING DATE: 20-NOV-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/031,435  
; FILING DATE: 20-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adler, Reid G.  
; REGISTRATION NUMBER: 30,988  
; REFERENCE/DOCKET NUMBER: 044574-5022-01-WO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-467-7000  
; TELEFAX: 202-467-7176  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 142 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-975-080-34

Query Match 100.0%; Score 778; DB 3; Length 142;  
Best Local Similarity 100.0%; Pred. No. 1e-79;  
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGAPTLPPAQWQPFKDHRISTFKNWFFLEGCACTPERMAEAGFIHCPTNEPDLAQCFFC 60  
DB 1 MGAPTLPPAQWQPFKDHRISTFKNWFFLEGCACTPERMAEAGFIHCPTNEPDLAQCFFC 60

QY 61 FKELEGWEPDDPIBEHKKHSGGCAFLSVKQFEELTLGFEFLKLDREPAKKNIAKETNNK 120  
 DB 61 FKELEGWEPDDPIBEHKKHSGGCAFLSVKQFEELTLGFEFLKLDREPAKKNIAKETNNK 120  
 QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 DB 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 2  
 US-09-283-144-3  
 ; Sequence 3, Application US/09283144  
 ; Patent No. 6346389  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yale University  
 ; TITLE OF INVENTION: Method for Selectively Modulating the Interactions  
 ; FILE REFERENCE: 44574-5033-US  
 ; CURRENT APPLICATION NUMBER: US/09/283,144  
 ; CURRENT FILING DATE: 1999-04-01  
 ; EARLIER APPLICATION NUMBER: US 60/080,288  
 ; EARLIER FILING DATE: 1998-04-01  
 ; EARLIER APPLICATION NUMBER: US 08/975,080  
 ; EARLIER FILING DATE: 1997-11-20  
 ; EARLIER APPLICATION NUMBER: PCT/US97/21880  
 ; EARLIER FILING DATE: 1997-11-20  
 ; EARLIER APPLICATION NUMBER: US 60/031,435  
 ; EARLIER FILING DATE: 1996-11-20  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 3  
 ; LENGTH: 142  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Amino acid sequence of Survivin  
 US-09-283-144-3

Query Match 100.0%; Score 778; DB 4; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 1e-79;  
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGAPTLPPAWQFLKDHRISTFKNWPFLGCACTPERMAEAGFIHCPTNEPDLAQCFFC 60  
 DB 1 MGAPTLPPAWQFLKDHRISTFKNWPFLGCACTPERMAEAGFIHCPTNEPDLAQCFFC 60  
 QY 61 FKELEGWEPDDPIBEHKKHSGGCAFLSVKQFEELTLGFEFLKLDREPAKKNIAKETNNK 120  
 DB 61 FKELEGWEPDDPIBEHKKHSGGCAFLSVKQFEELTLGFEFLKLDREPAKKNIAKETNNK 120  
 QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 DB 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 3  
 US-09-738-946-4  
 ; Sequence 4, Application US/09738946  
 ; Patent No. 6579701  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EXELIXIS, INC.  
 ; TITLE OF INVENTION: DROSOPHILA HOMOLOGUES OF GENES AND PROTEINS IMPLICATED IN CANCER  
 ; FILE REFERENCE: EX00-043C  
 ; CURRENT APPLICATION NUMBER: US/09/738,946  
 ; CURRENT FILING DATE: 2000-12-14  
 ; PRIOR APPLICATION NUMBER: 60/170,832  
 ; PRIOR FILING DATE: 1999-12-14  
 ; PRIOR APPLICATION NUMBER: 60/170,838  
 ; PRIOR FILING DATE: 1999-12-14  
 ; PRIOR APPLICATION NUMBER: 60/178,580  
 ; PRIOR FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: 60/185,879  
 ; PRIOR FILING DATE: 2000-02-29  
 ; PRIOR APPLICATION NUMBER: 60/185,880  
 ; PRIOR FILING DATE: 2000-02-29  
 ; PRIOR APPLICATION NUMBER: 60/186,150  
 ; PRIOR FILING DATE: 2000-03-01  
 ; PRIOR APPLICATION NUMBER: 60/189,701  
 ; PRIOR FILING DATE: 2000-03-15  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 153  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yale University  
 ; TITLE OF INVENTION: Method for Selectively Modulating the Interactions  
 ; FILE REFERENCE: 44574-5033-US  
 ; CURRENT APPLICATION NUMBER: US/09/283,144  
 ; CURRENT FILING DATE: 1999-04-01  
 ; EARLIER APPLICATION NUMBER: US 60/080,288  
 ; EARLIER FILING DATE: 1998-04-01  
 ; EARLIER APPLICATION NUMBER: US 08/975,080  
 ; EARLIER FILING DATE: 1997-11-20  
 ; EARLIER APPLICATION NUMBER: PCT/US97/21880  
 ; EARLIER FILING DATE: 1997-11-20  
 ; EARLIER APPLICATION NUMBER: US 60/031,435  
 ; EARLIER FILING DATE: 1996-11-20  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 15  
 ; LENGTH: 37  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Amino acid sequence from beginning of coding region of  
 ; OTHER INFORMATION: Survivin gene  
 US-09-283-144-15

Query Match 28.1%; Score 219; DB 4; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-17;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGAPTLPPAWQFLKDHRISTFKNWPFLGCACTPER 37  
 DB 1 MGAPTLPPAWQFLKDHRISTFKNWPFLGCACTPER 37  
 RESULT 5  
 US-08-836-134-23  
 ; Sequence 23, Application US/08836134A  
 ; Patent No. 6020127  
 ; GENERAL INFORMATION:

APPLICANT: Mackenzie, Alex E.  
APPLICANT: Korneluk, Robert G.  
APPLICANT: Mahadevan, Mani S.  
APPLICANT: McLean, Michael  
APPLICANT: Roy, Natalie  
APPLICANT: Ikeda, John-e  
TITLE OF INVENTION: Neuronal Apoptosis Inhibitor Protein, Gene Sequence and  
Patent No. 6020127  
TITLE OF INVENTION: Mutations Causative of Spinal Muscular Atrophy  
FILE REFERENCE: 3477-112, 033477/139914  
CURRENT APPLICATION NUMBER: US/08/836,134A  
CURRENT FILING DATE: 1997-06-20  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 23  
LENGTH: 1151  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-836-134-23

Query Match 19.9%; Score 155; DB 3; Length 1151;  
Best Local Similarity 40.5%; Pred. No. 1.4e-08;  
Matches 34; Conservative 15; Mismatches 29; Indels 6; Gaps 3;  
Qy 15 KDHRISTFKWPF-LEGCACTPERMAEAGFIHCPTENEPDLAQCFKCELEGWEPDDDP 73  
Db 168 EEARLASFRNWPFFVQG--ISPCVLSEAGFVF---TGKQDTVQCFCGCGCLGNWEEGDDP 222  
Qy 74 IEHKKHSSGCALSVKKQFEELT 97  
Db 223 WKEHAKWFPKCEFLRSKKSSEIT 246

RESULT 6  
US-09-493-784-23  
Sequence 23, Application US/09493784  
Patent No. 6429011  
GENERAL INFORMATION:  
APPLICANT: Mackenzie, Alex E.  
APPLICANT: Korneluk, Robert G.  
APPLICANT: Mahadevan, Mani S.  
APPLICANT: McLean, Michael  
APPLICANT: Roy, Natalie  
APPLICANT: Ikeda, John-e  
TITLE OF INVENTION: Neuronal Apoptosis Inhibitor Protein, Gene Sequence and  
Patent No. 6429011  
TITLE OF INVENTION: Mutations Causative of Spinal Muscular Atrophy  
FILE REFERENCE: 3477-112, 033477/139914  
CURRENT APPLICATION NUMBER: US/09/493,784  
CURRENT FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 08/836,134  
PRIOR FILING DATE: 1997-06-20  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 23  
LENGTH: 1151  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-493-784-23

Query Match 19.9%; Score 155; DB 4; Length 1151;  
Best Local Similarity 40.5%; Pred. No. 1.4e-08;  
Matches 34; Conservative 15; Mismatches 29; Indels 6; Gaps 3;  
Qy 15 KDHRISTFKWPF-LEGCACTPERMAEAGFIHCPTENEPDLAQCFKCELEGWEPDDDP 73  
Db 168 EEARLASFRNWPFFVQG--ISPCVLSEAGFVF---TGKQDTVQCFCGCGCLGNWEEGDDP 222  
Qy 74 IEHKKHSSGCALSVKKQFEELT 97  
Db 223 WKEHAKWFPKCEFLRSKKSSEIT 246

RESULT 7  
US-08-836-134-2  
Sequence 2, Application US/08836134A  
Patent No. 6020127  
GENERAL INFORMATION:  
APPLICANT: Mackenzie, Alex E.  
APPLICANT: Korneluk, Robert G.  
APPLICANT: Mahadevan, Mani S.  
APPLICANT: McLean, Michael  
APPLICANT: Roy, Natalie  
APPLICANT: Ikeda, John-e  
TITLE OF INVENTION: Neuronal Apoptosis Inhibitor Protein, Gene Sequence and  
Patent No. 6020127  
TITLE OF INVENTION: Mutations Causative of Spinal Muscular Atrophy  
FILE REFERENCE: 3477-112, 033477/139914  
CURRENT APPLICATION NUMBER: US/08/836,134A  
CURRENT FILING DATE: 1997-06-20  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1232  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-836-134-2

Query Match 19.9%; Score 155; DB 3; Length 1232;  
Best Local Similarity 40.5%; Pred. No. 1.5e-08;  
Matches 34; Conservative 15; Mismatches 29; Indels 6; Gaps 3;  
Qy 15 KDHRISTFKWPF-LEGCACTPERMAEAGFIHCPTENEPDLAQCFKCELEGWEPDDDP 73  
Db 159 EEARLASFRNWPFFVQG--ISPCVLSEAGFVF---TGKQDTVQCFCGCGCLGNWEEGDDP 213  
Qy 74 IEHKKHSSGCALSVKKQFEELT 97  
Db 214 WKEHAKWFPKCEFLRSKKSSEIT 237

RESULT 8  
US-09-493-784-2  
Sequence 2, Application US/09493784  
Patent No. 6429011  
GENERAL INFORMATION:  
APPLICANT: Mackenzie, Alex E.  
APPLICANT: Korneluk, Robert G.  
APPLICANT: Mahadevan, Mani S.  
APPLICANT: McLean, Michael  
APPLICANT: Roy, Natalie  
APPLICANT: Ikeda, John-e  
TITLE OF INVENTION: Neuronal Apoptosis Inhibitor Protein, Gene Sequence and  
Patent No. 6429011  
TITLE OF INVENTION: Mutations Causative of Spinal Muscular Atrophy  
FILE REFERENCE: 3477-112, 033477/139914  
CURRENT APPLICATION NUMBER: US/09/493,784  
CURRENT FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 08/836,134  
PRIOR FILING DATE: 1997-06-20  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1232  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-493-784-2

Query Match 19.9%; Score 155; DB 4; Length 1232;  
Best Local Similarity 40.5%; Pred. No. 1.5e-08;  
Matches 34; Conservative 15; Mismatches 29; Indels 6; Gaps 3;  
Qy 15 KDHRISTFKWPF-LEGCACTPERMAEAGFIHCPTENEPDLAQCFKCELEGWEPDDDP 73  
Db 159 EEARLASFRNWPFFVQG--ISPCVLSEAGFVF---TGKQDTVQCFCGCGCLGNWEEGDDP 213

QY 74 IEHKKHSSGCAFLSVKKQFELT 97  
Db 214 WKEHAKWPKCEFLRSKKSSEIT 237

RESULT 9  
US-09-705-872-3  
; Sequence 3, Application US/09705872  
; Patent No. 6617429  
; GENERAL INFORMATION:  
; APPLICANT: Joh-E IKEDA  
; APPLICANT: Kenji YAMAMOTO  
; TITLE OF INVENTION: APOPTOSIS INHIBITORY PROTEIN, GENE ENCODING THE PROTEIN  
; FILE REFERENCE: 2000-1110/LC/00653  
; CURRENT APPLICATION NUMBER: US/09/705,872  
; CURRENT FILING DATE: 2000-11-06  
; PRIOR APPLICATION NUMBER: 09/239,797  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1295  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-705-872-3

Query Match 19.9%; Score 155; DB 4; Length 1295;  
Best Local Similarity 40.5%; Pred. No. 1.6e-08;  
Matches 34; Conservative 15; Mismatches 29; Indels 6; Gaps 3;

QY 15 KDHRISTFKWPP-LEGCACTPERMAEAGFIHCTPTNEPDLAQCFPCFKEGWEPPDDP 73  
Db 159 BEARLASFRNPFYVQG--ISPCVLSEAGFV---TGKQDTVQCFSCGCLGNWEEGDDP 213  
QY 74 IEHKKHSSGCAFLSVKKQFELT 97  
Db 214 WKEHAKWPKCEFLRSKKSSEIT 237

RESULT 10  
US-09-705-872-1  
; Sequence 1, Application US/09705872  
; Patent No. 6617429  
; GENERAL INFORMATION:  
; APPLICANT: Joh-E IKEDA  
; APPLICANT: Kenji YAMAMOTO  
; TITLE OF INVENTION: APOPTOSIS INHIBITORY PROTEIN, GENE ENCODING THE PROTEIN  
; FILE REFERENCE: 2000-1110/LC/00653  
; CURRENT APPLICATION NUMBER: US/09/705,872  
; CURRENT FILING DATE: 2000-11-06  
; PRIOR APPLICATION NUMBER: 09/239,797  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1403  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-705-872-1

Query Match 19.9%; Score 155; DB 4; Length 1403;  
Best Local Similarity 40.5%; Pred. No. 1.8e-08;  
Matches 34; Conservative 15; Mismatches 29; Indels 6; Gaps 3;

QY 15 KDHRISTFKWPP-LEGCACTPERMAEAGFIHCTPTNEPDLAQCFPCFKEGWEPPDDP 73  
Db 159 BEARLASFRNPFYVQG--ISPCVLSEAGFV---TGKQDTVQCFSCGCLGNWEEGDDP 213  
QY 74 IEHKKHSSGCAFLSVKKQFELT 97  
Db 214 WKEHAKWPKCEFLRSKKSSEIT 237

RESULT 11  
US-08-657-759-22  
; Sequence 22, Application US/08657759  
; Patent No. 6511828  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Craig B.  
; APPLICANT: Duckett, Colin S.  
; TITLE OF INVENTION: HUMAN AND DROSOPHILA INHIBITORS  
; TITLE OF INVENTION: OF APOPTOSIS PROTEINS (IAPs)  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P. O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: United States of America  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/657,759  
; APPLICATION NUMBER: US/08/657,759  
; FILING DATE: 31-MAY-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Highlander, Steven L.  
; REGISTRATION NUMBER: 37,642  
; REFERENCE/DOCKET NUMBER: ARCD:220  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (512) 474-7577  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 75 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-657-759-22

Query Match 18.0%; Score 140; DB 4; Length 75;  
Best Local Similarity 39.7%; Pred. No. 2e-08;  
Matches 31; Conservative 14; Mismatches 27; Indels 6; Gaps 3;

QY 15 KDHRISTFKWPP-LEGCACTPERMAEAGFIHCTPTNEPDLAQCFPCFKEGWEPPDDP 73  
Db 3 BEARLASFRNPFYVQG--ISPCVLSEAGFV---TGKQDTVQCFSCGCLGNWEEGDDP 57  
QY 74 IEHKKHSSGCAFLSVKK 91  
Db 58 WKEHAKWPKCEFLRSK 75

RESULT 12  
US-08-511-485-12  
; Sequence 12, Application US/08511485  
; Patent No. 5919912  
; GENERAL INFORMATION:  
; APPLICANT: Korneluk, Robert G.  
; APPLICANT: Mackenzie, Alexander E.  
; APPLICANT: Baird, Stephen  
; TITLE OF INVENTION: MAMMALIAN IAP GENE FAMILY, PRIMERS,  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA

```
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/511,485
; FILING DATE: 04-AUG-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 07540/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: both
; MOLECULE TYPE: protein
; US-08-511-485-12

Query Match 17.5%; Score 136.5; DB 2; Length 275;
Best Local Similarity 36.4%; Pred. No. 2.6e-07;
Matches 28; Conservative 14; Mismatches 24; Indels 11; Gaps 3;

QY 18 RISTKKNWPLEGCAC---TPERMAEAGFIHCPTENEPDLAQCFCFKELSGWEPDDDDPI 74
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
111 RVKSFHNP-----RCMKORPEQMADAGFFY---TGYGNTKCFYCDGGLKDWEPEDVPW 162
QY 75 EHKHSSGCAFLSVKX 91
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
163 EQHVRWFDRCAVQLVK 179

RESULT 13
US-08-836-134-21
; Sequence 21, Application US/08836134A
; Patent No. 6020127
; GENERAL INFORMATION:
; APPLICANT: Mackenzie, Alex E.
; APPLICANT: Korneluk, Robert G.
; APPLICANT: Mahadevan, Mani S.
; APPLICANT: McLean, Michael
; APPLICANT: Roy, Natalie
; APPLICANT: Ikeda, John-e
; TITLE OF INVENTION: Neuronal Apoptosis Inhibitor Protein, Gene Sequence and
; FILE REFERENCE: 3477-112, 033477/139914
; CURRENT APPLICATION NUMBER: US/08/836,134A
; CURRENT FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Cydia pomonella
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PCR primer
US-08-836-134-21

Query Match 17.5%; Score 136.5; DB 4; Length 275;
Best Local Similarity 36.4%; Pred. No. 2.6e-07;
Matches 28; Conservative 14; Mismatches 24; Indels 11; Gaps 3;

QY 18 RISTKKNWPLEGCAC---TPERMAEAGFIHCPTENEPDLAQCFCFKELSGWEPDDDDPI 74
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
111 RVKSFHNP-----RCMKORPEQMADAGFFY---TGYGNTKCFYCDGGLKDWEPEDVPW 162
QY 75 EHKHSSGCAFLSVKX 91
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
163 EQHVRWFDRCAVQLVK 179

RESULT 14
US-09-493-784-21
; Sequence 21, Application US/09493784
; Patent No. 6429011
; GENERAL INFORMATION:
; APPLICANT: Mackenzie, Alex E.
; APPLICANT: Korneluk, Robert G.
; APPLICANT: Mahadevan, Mani S.
; APPLICANT: McLean, Michael
; APPLICANT: Roy, Natalie
; APPLICANT: Ikeda, John-e
; TITLE OF INVENTION: Neuronal Apoptosis Inhibitor Protein, Gene Sequence and
; FILE REFERENCE: 3477-112, 033477/139914
; CURRENT APPLICATION NUMBER: US/09/493,784
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 08/836,134
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Cydia pomonella
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PCR primer
US-09-493-784-21

Query Match 17.5%; Score 136.5; DB 4; Length 275;
Best Local Similarity 36.4%; Pred. No. 2.6e-07;
Matches 28; Conservative 14; Mismatches 24; Indels 11; Gaps 3;

QY 18 RISTKKNWPLEGCAC---TPERMAEAGFIHCPTENEPDLAQCFCFKELSGWEPDDDDPI 74
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
111 RVKSFHNP-----RCMKORPEQMADAGFFY---TGYGNTKCFYCDGGLKDWEPEDVPW 162
QY 75 EHKHSSGCAFLSVKX 91
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
163 EQHVRWFDRCAVQLVK 179

RESULT 15
US-09-201-936-12
; Sequence 12, Application US/09201936
; Patent No. 6541457
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G.
; APPLICANT: Mackenzie, Alexander E.
; APPLICANT: Baird, Stephen
; APPLICANT: Liston, Peter
; TITLE OF INVENTION: MAMMALIAN IAP GENE FAMILY, PRIMERS,
; FILE REFERENCE: 07891/003003
; CURRENT APPLICATION NUMBER: US/09/201,936
; CURRENT FILING DATE: 1998-12-01
; EARLIER APPLICATION NUMBER: 09/011,356
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: PCT/IB96/01022
; EARLIER FILING DATE: 1996-08-05
; EARLIER APPLICATION NUMBER: 08/576,956
; EARLIER FILING DATE: 1995-12-22
; EARLIER APPLICATION NUMBER: 08/511,485
; EARLIER FILING DATE: 1995-08-04
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 12
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Cydia pomonella
US-09-201-936-12

Query Match      17.5%; Score 136.5; DB 4; Length 275;
Best Local Similarity 36.4%; Pred. No. 2.6e-07;
Matches 28; Conservative 14; Mismatches 24; Indels 11; Gaps 3;

QY  18 RISTFQWPFLEGCAC---TPERMAEAGFIHCPTENEPDLAQCFFCFKELEGWEPDDPI 74
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  111 RVKSFHNWP-----RCMKQPEQWADAGFFY--TGyDNTKCFYCDGGLKDWEPEdVPW 162
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  75 BEHKHSHSGCAFLSVKK 91
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  163 EQHVRWFDRCAYVOLVK 179
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: August 11, 2004, 14:18:05  
Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August '15, 2004, 21:13:29 ; Search time 2786 Seconds  
(without alignments)  
1522.049 Million cell updates/sec

Title: US-09-690-825-34  
Perfect score: 778  
Sequence: 1 MGAPLPPAWQFLKDHRS.....EFETAKKVRRAIEQLAAMD 142

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+pn.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US09690825/runat\_11082004\_140952\_13543/app\_query.fasta\_1.327  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09690825 @CGN 1.1 3437 @runat\_11082004\_140952\_13543 -NCPUs=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_htc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estcom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_man.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_pig.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	778	100.0	472	12	BM754864
2	778	100.0	478	9	AL710957
3	778	100.0	517	10	AW411195
4	778	100.0	552	12	BG258128
5	778	100.0	552	12	BM839824
6	778	100.0	553	14	CB138359
7	778	100.0	563	12	BM835829
8	778	100.0	574	14	CB111125
9	778	100.0	580	10	AW409701
10	778	100.0	584	14	CB128010
11	778	100.0	585	12	BG776624
12	778	100.0	602	9	AL603004
13	778	100.0	605	12	BM841904
14	778	100.0	613	10	AW409816
15	778	100.0	624	14	CB135918
16	778	100.0	629	10	BE292733
17	778	100.0	644	12	BM788077
18	778	100.0	646	12	B1252508
19	778	100.0	682	10	BE796084
20	778	100.0	698	10	BE796061
21	778	100.0	703	10	BF219993
22	778	100.0	707	12	BG115707
23	778	100.0	718	12	BM827836
24	778	100.0	726	13	BX421971
25	778	100.0	742	10	BE907403
26	778	100.0	760	12	BM020628
27	778	100.0	766	12	BG682887
28	778	100.0	768	12	B1771953
29	778	100.0	772	12	BG750227
30	778	100.0	785	13	BU600185
31	778	100.0	815	12	BG335442
32	778	100.0	815	12	BG744230
33	778	100.0	821	12	BM542207
34	778	100.0	822	9	AU130493
35	778	100.0	826	13	BU171608
36	778	100.0	828	13	BU600854
37	778	100.0	836	9	AU126048
38	778	100.0	861	13	BQ438198
39	778	100.0	868	13	BQ832604
40	778	100.0	874	10	BE883968
41	778	100.0	875	10	BE898710
42	778	100.0	875	12	BG528273
43	778	100.0	880	12	BM450871
44	778	100.0	883	13	BQ650551
45	778	100.0	888	13	BQ227378

# ALIGNMENTS

RESULT 1  
BM754864  
LOCUS K-EST0032467 S11SNUI Homo sapiens cDNA clone S11SNUI-12-C02 5',  
DEFINITION 472 bp mRNA linear EST 04-MAR-2002  
ACCESSION BM754864  
VERSION BM754864.1 GI:19084482  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 472)

**AUTHORS** Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.

**TITLE** 21C Frontier Korean EST Project 2001

**JOURNAL**

**COMMENT**

Genome Research Center

Contact: Kim YS

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.krribb.re.kr

Plate: 12 row: C column: 02

High quality sequence, scop: 472.

**FEATURES**

source

Location/Qualifiers

1..472

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S11SNU1-12-C02"

/sex="W"

/tissue\_type="Stomach"

/cell\_type="Lymphoblast-like"

/cell\_line="SNU-1"

/lab\_host="Top10F"

/clone\_lib="S11SNU1"

/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;

Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then deacapped  
with tabacco acid pyrophosphatase (TAP). The deacapped  
intact mRNA was ligated with DNA-RNA linker including SfiI  
site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized with Superscript II using SfiI  
oligo-3T primer. After first strand synthesis, RNA was  
degraded by NaOH treatment and cDNA was amplified by PCR  
reaction. The PCR products were digested with SfiI and  
cloned into pT7-1 digested pME18-FL3 vector. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."

**ORIGIN**

Alignment Scores:

Pred. No.: 2,61e-74 Length: 472  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-690-825-34 (1-142) x BM754864 (1-472)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 45 ATGGTGCCCGACGTTGGCCCTTCTGAGAGCCAGCTTCTCAAGGACACCCGATCTCT 104  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 105 ACATTCAAGAACTGGCCCTCTTGGAGGGCTGGCCCTGACCCCGAGGATGGCCGAG 164  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 165 GCTGGCTTCATCACTCCCTCCCTAGAGAGCCAGACTTGGCCCGAGTTCCTCTCTGC 224  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80  
DB 225 TTCAAGAGAGCTGGAAGCTGGAGCCAGATGACGCCCATATAGGACATATAAAGCAT 284  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 285 TCGTCCGGTTGCGCTTTCCTTCTGTCAAGAGCAGTTTGAAGATTAACCCCTTGTGAA 344

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
DB 345 TTTTGAAGAACTGCAGAGAAAGACCAAGCAAAATTTGCAAGGAAACCAACAATAAG 404  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaLa 140  
DB 405 AAGAAGAATTTGAGGAACTGCGAAGAAAGTGCCCGTGCATCGAGCAGCTGGCTGCC 464  
QY 141 MetAsp 142  
DB 465 ATGGAT 470  
RESULT 2  
AL710957 478 bp mRNA linear EST 04-SEP-2003  
LOCUS DKFZp686A0177.r1.686 (synonym: hlcc3) Homo sapiens cDNA clone  
DEFINITION DKFZp686A0177.5', mRNA sequence.  
ACCESSION AL710957  
VERSION AL710957  
KEYWORDS EST.  
SOURCE Homo sapiens (human).  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 478)  
AUTHORS Wambutt,R., Heubner,D., Mewes,W., Weil,B. and Wiemann,S.  
TITLE EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B. and Wiemann,S.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: MIPS  
MIPS  
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by AGOVA (Berlin/Germany) within the cDNA sequencing  
consortium of the German Genome Project.  
No 5' sequence available.  
This clone (DKFZp686A0177) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
FEATURES  
Location/Qualifiers  
1..478  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686A0177"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="686 (synonym: hlcc3)"  
/note="Vector: pTriplex2; Site\_1: SfiIA; Site\_2: SfiIB;  
cDNA-collection"

**ORIGIN**

Alignment Scores:  
Pred. No.: 2,65e-74 Length: 478  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x AL710957 (1-478)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 22 ATGGTGCCCGACGTTGGCCCTTCTGAGAGCCAGCTTCTCAAGGACACCCGATCTCT 81  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 82 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGGCCCTGACCCCGAGGCGGATGGCCGAG 141  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60



```

Db      142 GCTGCTTCATCCACTGCCCACTGAGAACGAGCAGACTTGGCCCGAGTGTTCCTTCGC 201
QY      61 PheLysGluLeuGluGluGluTrpGluProAspAspAspProIleGluHisLysLysHis 80
Db      202 TTCAGGAGCTGGAAGGCTGGAGCGACATCAGCAGCCCATAGAGGAACATAAAAGCAT 261
QY      81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGluGlu 100
Db      262 TCGTCCGTTGGCTTCTCTCTGTCAGAGCAGAGTTGAGGAATTAACCTTGGTGAA 321
QY      101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLysLysLysLysLysLys 120
Db      322 TTTTGAACCTGGACAGAGAGAGCCAGCAACAAATTCAGAGGAACCAACAATAAG 381
QY      121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaLysLysLysLysLysLys 140
Db      382 AAGAAAGAAATTTAGGAAACTGCGAAGAAAGTGGCGCGTGCATGCGAGCAGCTGGCTGCC 441
QY      141 MetAsp 142
Db      442 ATGGAT 447

RESULT 3
AW411195
LOCUS   f10h09_x1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2964713 5',
DEFINITION mRNA sequence.
ACCESSION AW411195
VERSION   AW411195.1 GI:6936736
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 517)
          NIH-MGC http://mgi.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Ling Hong/Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: National Institutes of Health Intramural
          Sequencing Center (NISC)
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          www.bio.llnl.gov/bbrp/image/image.html
          Plate: LCM64 row: C column: 18
          Seq primer: -21M13 forward primer (ABI).
          Location/Qualifiers
            1..517
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:2964713"
              /tissue_type="thadomyosarcoma"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH MGC 17"
              /notes="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;
              Site 2: XhoI; cDNA made by oligo-dT priming.
              Directionally cloned into EcoRI/XhoI sites using the
              following 5' adaptor: GGACGAG(G). Size-selected >500bp
              for average insert size 1.8kb. Library constructed by
              Ling Hong in the laboratory of Gerald M. Rubin (University
              of California, Berkeley) using ZAP-cDNA synthesis kit
              (Stratagene) and Superscript II RT (Life Technologies)."
```

## ORIGIN

```

Alignment Scores:      2,91e-74      Length:      517
Pred. No.:            778.00      Matches:      142
Score:                100.00%      Conservative: 0
Percent Similarity:
```

```

Best Local Similarity: 100.00%      Mismatches: 0
Query Match:          100.00%      Indels: 0
DB:                  10           Gaps: 0
US-09-690-825-34 (1-142) x AW411195 (1-517)

QY      1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db      44 ATGGTGTCCCGGACGTTGCCCTGCTGCGAGCCCTTTCTCAAGGACCACCGCATCTCT 103
QY      21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db      104 ACATTCAAGAACTGGCCCTTTCTTGGAGGGCTGCGCTGCACCCCGAGCGGATGGCGGAG 163
QY      41 AlaGlyPheIleHisCysPProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db      164 GCTGGCTTTCATCCATGCCCTGAGACGACGACGACTTGGCCCGAGTGTTCCTTCGTC 223
QY      61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysLysHis 80
Db      224 TTCAAGGAGCTGGAAGCTGGGAGCCAGATGACGACCCCATAGAGGAACATAAAAGCAT 283
QY      81 SerSerGlyCysAlaPheLeuSerValLysLysLysLysLysLysLysLysLysLys 100
Db      284 TCGTCCGTTGGCTTCTCTCTGTCAGAGACAGTTTGAAGAATTAACCTTGGTGAA 343
QY      101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLysLysLysLysLysLys 120
Db      344 TTTTGAACCTGGACAGAGAGAGCAACAAATTCGAAAGGAACCAACAATAAG 403
QY      121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaLysLysLysLysLys 140
Db      404 AAGAAAGAAATTTAGGAAACTGCGAAGAAAGTGGCGCGTGCATGCGAGCAGCTGGCTGCC 463
QY      141 MetAsp 142
Db      464 ATGGAT 469

RESULT 4
BG258128
LOCUS   602379226F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4510014 5',
DEFINITION mRNA sequence.
ACCESSION BG258128
VERSION   BG258128.1 GI:12767944
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 552)
          NIH-MGC http://mgi.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLM0391 row: C column: 07
          High quality sequence stop: 552.
          Location/Qualifiers
            1..552
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:4510014"
              /tissue_type="embryonal carcinoma, cell line"
              /lab_host="DH10B (phage-resistant)"
```

## FEATURES

## source

/clone.lib="NIH\_MGC\_92"  
 /note="Organ: testis; Vector: PCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 2.5 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,14e-74 Length: 552  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservativeness: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-09-690-825-34 (1-142) x BG258128 (1-552)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 DB 32 ATGGTGGCCCGAGCTTGCCCTTGCCTGGGAGCCCTTCTCAAGGACCCACCATCTCT 91  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 92 ACATTCAAGAACTGGCCCTTCTGGAGGCTGGCCTGCACCCCGAGCGATGGCCGAG 151  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
 DB 152 GCTGGCTTCATCCACTGCCCTCCCTGAGACGAGCAGACTGGCCAGTGTTCCTCTGC 211  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
 DB 212 TTCAGAGAGCTGGAGGCTGGAGCCAGATGACGACCCCATAGAGGACATATAAGCAT 271  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 DB 272 TCGTCCGGTTCGGCTTTCCTTCTGTCAAGAGCAGTTTCAAGAAATTAACCTTGGTGA 331  
 QY 101 PheLeuLysLeuAspArgGluArgCysAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 DB 332 TTTTGAACATGGACAGAAAGAGCCCAAGAAACAAATTCAGAGGAAACCAACAATAG 391  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
 DB 392 AAGAAAGAAATTTGAGGAAATCGGAGAAAGTGGCCGTCCTCATCGAGCAGCTGGCTGC 451  
 QY 141 MetAsp 142  
 DB 452 ATGGAT 457

RESULT 5  
 BM839824  
 LOCUS  
 DEFINITION K-ESR0116794 S13KM55 Homo sapiens cDNA clone S13KM55-32-F06 5',  
 mRNA sequence.  
 ACCESSION BM839824  
 VERSION BM839824.1 GI:19196233  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1. (bases 1 to 552)  
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470

Fax: +82-42-860-4409  
 Email: yongsung@mail.kr.ibm.re.kr  
 Plate: 32 row; F column: 06  
 High quality sequence stop: 552.  
 Location/Qualifiers  
 1. .552

FEATURES  
 source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S13KM55-32-F06"  
 /tissue\_type="myeloma"  
 /cell\_line="KMS-5"  
 /lab\_host="Top10F"  
 /clone\_lib="S13KM55"  
 /note="Vector: pcns; Site 1: EcoRI; Site 2: NotI; The poly  
 (A) + RNA was dephosphorylated with bacterial alkaline  
 phosphatase (BAP) and then decapped with tobacco acid  
 pyrophosphatase (TAP). The decapped intact mRNA was  
 ligated with DNA-RNA linker including EcoR I site by  
 treatment of T4 RNA ligase and the first strand cDNA was  
 synthesized from oligo dt-selected mRNA by priming with  
 dt-tailed vector. The dt-tailed vector was adjusted to  
 have about 60nt. The cDNA vector was circularized with E.  
 coli DNA ligase after digestion of EcoRI which site is  
 also included in vector. An RNA strand converted to a DNA  
 strand by Okayama-Berg method. The obtained cDNA vectors  
 were used for transformation of competent cells E. coli  
 Top10F by electroporation method. The cDNA libraries  
 constructed by this method are full-length enriched cDNA  
 library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,14e-74 Length: 552  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservativeness: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-690-825-34 (1-142) x BM839824 (1-552)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 DB 78 ATGGTGGCCCGAGCTTGCCCTTGCCTGGCAGCCCTTCTCAAGGACCCACCATCTCT 137  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 138 ACATTCAAGAACTGGCCCTTCTGGAGGCTGGCCTGCACCCCGAGCGATGGCCGAG 197  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
 DB 198 CTTGGCTTCATCCACTGCCCTCCCTGAGACGAGCAGACTTGGCCCACTGTTCTTCTGC 257  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
 DB 258 TTCAAGAGAGCTGGAAGGCTGGAGCCAGATGACGACCCCATAGAGGAACATAAAGCAT 317  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluLeuThrLeuGlyGlu 100  
 DB 318 TCGTCCGGTTCGGCTTTCCTTCTGTCAAGAGCAGTTTGAAGAATTAACCTTGGTGAA 377  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 DB 378 TTTTGAACATGGACAGAAAGAGCCCAAGAAACAAATTCAGAGGAAACCAACAATAG 437  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
 DB 438 AAGAAAGAAATTTGAGGAAATCGGAGAAAGTGGCCGTCCTCATCGAGCAGCTGGCTGCC 497  
 QY 141 MetAsp 142  
 DB 498 ATGGAT 503

```

RESULT 6
CBI38359
LOCUS
DEFINITION K-EST0191218 L5HLK1 Homo sapiens cDNA clone L5HLK1-23-G07 5', mRNA
sequence.
ACCESSION CBI38359
VERSION CBI38359.1 GI:28109974
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 553)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 23 row: G column: 07
High quality sequence stop: 553.
FEATURES
Source
1..553
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L5HLK1-23-G07"
/sex="M"
/cell_line="HLK-1"
/lab_host="Top10p"
/clone_lib="L5HLK1"
/note="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10p by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
ORIGIN
Alignment Scores:
Pred. No.: 3.15e-74 Length: 553
Score: 778.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-09-690-825-34 (1-142) x CBI38359 (1-553)
Qy 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 40 ATGGGTGCCCCGAGGTTCGCCCTTCCTGGCAGCCCTTTCTCAAGGACCACCGCATCTCT 99
Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetIleGlu 40
Db 100 ACATTCAAGAACTGGCCCTTCTTGGAGGGGTGGCGCTGCACCCCGAGCGATGCCGAG 159

```

```

Qy 41 AlaGlyPheIleHisCysProThrGluAenGluProAspLeuAlaGlnCysPheCys 60
Db 160 GCTGGGTTTCATCCACATGCCCACTGAGAAAGAGCCAGACTTGGCCAGTGTTCCTTCGC 219
Qy 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisLysLysHis 80
Db 220 TTCAGGAGCTGGAGGCTGGGAGCGGATGAGCGACCCCATAGAGGACATATAAAGAT 279
Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
Db 280 TCGTCCGGTTCCTTCCTTCTGTCAAGAGAGCAGTTGAAGAATTAACCCCTTGGTGA 339
Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLysLysLysLysLysLys 120
Db 340 TTTTGAAGTGGACAGAGAGAGGACCAACACCAAAATTCGCAAGGAAACCAACATAG 399
Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAla 140
Db 400 AAGAAAGAAATTTGAGGAAACTGCAAGAAAGTGGCGCGTGCATCGAGCAGCTGCC 459
Qy 141 MetAsp 142
Db 460 ATGGAT 465
RESULT 7
CBI38359
LOCUS
DEFINITION K-EST0111069 S11SNUI Homo sapiens cDNA clone S11SNUI-73-H09 5',
mRNA sequence.
ACCESSION CBI38359
VERSION CBI38359.1 GI:19192238
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 563)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 73 row: H column: 09
High quality sequence stop: 563.
FEATURES
Source
1..563
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S11SNUI-73-H09"
/sex="M"
/tissue_type="Stomach"
/cell_type="Lymphoblast-like"
/cell_line="SNU-1"
/lab_host="Top10p"
/clone_lib="S11SNUI"
/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
Site 2: XhoI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR

```

reaction. The PCR products were digested with SfiI and cloned into DraIII-digested pME18S-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells *E. coli* Top10<sup>®</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3.22e-74 Length: 563  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservativity: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-09-690-825-34 (1-142) x BM835829 (1-563)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 45 ATGGTGGCCGCGAGCTTGGAGGGGTGGCTGCGAGCCCTTTCTCAAGGACCAACCGCATCTCT 104  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 105 ACATTCAAGAACTGGCCCTTCTGGAGGGGTGGCTGCGAGCCCTTTCTCAAGGACCAACCGCATCTCT 164  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
 Db 165 GCTGGCTTCATCCACTGCCCTGAGACGAGCAGACTGGCCCGAGTTTCTTCTGTC 224  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80  
 Db 225 TTCAGAGAGCTGGAGCTGGAGCCAGATGACGACCCCATAGAGGAACATAAAGCAT 284  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 285 TCGTCCGGTTGCGCTTCTCTTCTGTCAGAACGAGCTTTGAAGAATTAACCTTGGTGA 344  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 Db 345 TTTTGAACCTGACAGAGAGAGAGCAAGAAATTTGCAAGAGAAACCAACATAG 404  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaLa 140  
 Db 405 AAGAAAGAAATTTGAGGAACTGCGAAGAGAGTGGCGCTGCCATCGAGCAGCTGCGTGC 464  
 QY 141 MetAsp 142  
 Db 465 ATGGAT 470

RESULT 8  
 CB111125  
 LOCUS  
 DEFINITION K-EST0152661 L5HLK1 Homo sapiens cDNA clone L5HLK1-8-G05 5', mRNA  
 sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470

Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 8 row; G column: 05  
 High quality sequence stop: 574.  
 Location/Qualifiers  
 1..574

FEATURES  
 source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="L5HLK1-8-G05"  
 /sex="M"  
 /cell\_line="HLK-1"  
 /lab\_host="Top10<sup>®</sup>"  
 /clone\_lib="L5HLK1"

/note="Organ: Liver; Vector: pcNS-D2; Site 1: EcoRI;  
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tobacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including  
 EcoRI site by treatment of T4 RNA ligase and the first  
 strand cDNA was synthesized from oligo dT-selected mRNA by  
 priming with dT-tailed vector. The dT-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was  
 circularized with *E. coli* DNA ligase after digestion of  
 EcoRI which site is also included in vector. An RNA strand  
 converted to a DNA strand by Okayama-Berg method. The  
 obtained cDNA vectors were used for transformation of  
 competent cells *E. coli* Top10<sup>®</sup> by electroporation method.  
 The cDNA libraries constructed by this method are  
 full-length enriched cDNA library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3.29e-74 Length: 574  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservativity: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 14 Gaps: 0

US-09-690-825-34 (1-142) x CB111125 (1-574)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 55 ATGGTGGCCGCGAGCTTGGAGGGGTGGCTGCGAGCCCTTTCTCAAGGACCAACCGCATCTCT 114  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 115 ACATTCAAGAACTGGCCCTTCTGGAGGGGTGGCTGCGAGCCCTTTCTCAAGGACCAACCGCATCTCT 174  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
 Db 175 GCTGGCTTCATCCACTGCCCTGAGACGAGCAGACTGGCCCGAGTTTCTTCTGTC 234  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80  
 Db 235 TTCAGAGAGCTGGAGGCTGGAGCCAGATGACGACCCCATAGAGGAACATAAAGCAT 294  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 295 TCGTCCGGTTGCGCTTCTCTTCTGTCAGAACGAGCTTTGAAGAATTAACCTTGGTGA 354  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 Db 355 TTTTGAACCTGACAGAGAGAGCAAGAAATTTGCAAGAGAAACCAACCAATAG 414  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaLa 140  
 Db 415 AAGAAAGAAATTTGAGGAACTGCGAAGAGAGTGGCGCTGCCATCGAGCAGCTGCGTGC 474  
 QY 141 MetAsp 142  
 Db 475 ATGGAT 480

```

RESULT 9
AW409701
LOCUS
DEFINITION
  fho1e02.x1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2961098 5',
  mRNA sequence.
ACCESSION
  AW409701
VERSION
  AW409701.1 GI:6935175
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 580)
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: ATCC
  cDNA Library Preparation: Ling Hong/Rubin Laboratory
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: National Institutes of Health Intramural
  Sequencing Center (NISC)
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.lnl.gov/bbrp/image/image.html
  Plate: LLCW55 row: I column: 3
  Seq primer: -21M13 forward primer (ABI).
FEATURES
  source
    1..580
    Location/Qualifiers
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:2961098"
      /tissue_type="rhadomyosarcoma"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH MGC 17"
      /note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;
      Site 2: XhoI; cDNA made by oligo-dT priming.
      Directionally cloned into EcoRI/XhoI sites using the
      following 5' adaptor: GGACGAG(G). Size-selected >500bp
      for average insert size 1.8kb. Library constructed by
      Ling Hong in the laboratory of Gerald M. Rubin (University
      of California, Berkeley) using ZAP-cDNA synthesis kit
      (Stratagene) and Superscript II RT (Life Technologies)."
```

## ORIGIN

```

Alignment Scores:
Pred. No.:      3,33e-74      Length:      580
Score:          778.00      Matches:     142
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches:  0
Query Match:     100.00%      Indels:      0
DB:              10          Gaps:          0
```

US-09-690-825-34 (1-142) x AW409701 (1-580)

```

Qy      1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db      53 ATGGGTGCCCCGACGTGTGCCCCCTGCTGCGACGCCCTTTCTCAGGACCCACCGCATCTCT 112
Qy      21 ThrPheLysAsnTrpProPheLeuGluGlyCyAlaCysThrProGluArgMetAlaGlu 40
Db      113 ACATTCAAGAACTGGCCCTCTTTGAGGGCTGCGCCTGCACCCCGAGCGGATGCCGAG 172
Qy      41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db      173 GCTGGGTTTCATCCAGTCCGCCCACTGAGACGACGACGACTTGCCCGAGGTTTCTTCGTGC 232
Qy      61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80
Db      233 TTCAGGAGCTGGAAGGCTGGGAGCGCAGATGACGACCCCATAGAGGAACATAAAGCAT 292
```

```

Qy      81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
Db      293 TCGTCGGGTTGCGCTTCTCTTGTCAAGAGACAGTTTGAAGATTAAACCCCTGGTGAA 352
Qy      101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
Db      353 TTTTGAACCTGCAGAGAAAGAGCGCAAAATTGCAAGAGGAACCAACAATAG 412
Qy      121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaLa 140
Db      413 AAGAAAGATTTCAGGAAACTGCGAAGAAAGTGCGCCGTGCCATCGAGCAGCTGGTGCC 472
Qy      141 MetAsp 142
Db      473 ATGGAT 478
RESULT 10
CB128010      584 bp      mRNA      linear      EST 29-JAN-2003
LOCUS
DEFINITION
  K-ES01017364 C1SNU17 Homo sapiens cDNA clone C1SNU17-12-G02 5',
  mRNA sequence.
ACCESSION
  CB128010
VERSION
  CB128010.1 GI:28090807
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 584)
  Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
  Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
  Kim,Y.S.
  21C Frontier Korean EST Project 2001
  Unpublished (2002)
  Contact: Kim YS
  Genome Research Center
  Korea Research Institute of Bioscience & Biotechnology
  52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
  Tel.: +82-42-860-4470
  Fax: +82-42-860-4409
  Email: yongsung@mail.kribb.re.kr
  Plate: 12 row: G column: 02
  High quality sequence stop: 584.
  Location/Qualifiers
    1..584
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="C1SNU17-12-G02"
    /sex="F"
    /tissue_type="Uterine"
    /cell_type="Epithelial"
    /cell_line="SNU-17"
    /lab_host="Top10F"
    /clone_lib="C1SNU17"
    /note="Organ: Cervix; Vector: pCNS-D2; Site 1: EcoRI;
    Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
    bacterial alkaline phosphatase (BAP) and then decapped
    with tabacco acid pyrophosphatase (TAP). The decapped
    intact mRNA was ligated with DNA-RNA linker including
    EcoRI site by treatment of T4 RNA ligase and the first
    strand cDNA was synthesized from oligo dt-selected mRNA by
    priming with dt-tailed vector. The dt-tailed vector was
    adjusted to have about 60nt. The cDNA vector was
    circularized with E. coli DNA ligase after digestion of
    EcoRI which site is also included in vector. An RNA strand
    converted to a DNA strand by Okayama-Berg method. The
    obtained cDNA vectors were used for transformation of
    competent cells E. coli Top10F by electroporation method.
    The cDNA libraries constructed by this method are
    full-length enriched cDNA library."
```

## ORIGIN

```

Alignment Scores:
Pred. NO.: 3.36e-74 Length: 584
Score: 778.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-690-825-34 (1-142) x CB128010 (1-584)

Qy 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 55 ATGGGTGCCCCGAGTTGGCCCTGCTGCTGAGCCCTTCTCAAGGACCCGATCTCT 114
Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 115 ACATTCAGAACTGGCCCTTCTTGGAGGCTGGCCCTGCACCCCGAGCGGATGGCCGAG 174
Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60
Db 175 GCTGGCTTCATCACTCCCTCCCATGTGAACAGCCAGACTTGGCCAGTGTCTTCTGTC 234
Qy 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80
Db 235 TTCAAGGAGCTGGAAGCTGGAGCTGGAGCCAGATGACGACCCCATAGAGGAACATAAAAGCAT 294
Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100
Db 295 TCGTCCCGTTGCGCTTCTTCTGTCGAAAGACAGTTTGAAGAATTAACCTTGTGTAA 354
Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
Db 355 TTTTGAACCTGGACAGAGAAAGAGCCAGCAACAAAAATTCGAAAGGAAACCAACATAAG 414
Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaA 140
Db 415 AAGAAAGAAATTTGAGGAATTCGGAAGAAAGTGGCCGTCCTGATCGAGCAGCTGGCTGCC 474
Qy 141 MetAsp 142
Db 475 ATGGAT 480

RESULT 11
BG776624 602663774F1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4808802 5',
LOCUS mRNA sequence.
ACCESSION BG776624
VERSION BG776624.1 GI:14046941
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://www.ncbi.nlm.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LCM:662 row: d column: 19
High quality sequence stop: 585.
Location/Qualifiers
1..585
/organism="Homo sapiens"

/mot_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4808802"
/tissue_type="mucoepidermoid carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_59"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgctcgcc); Site_2: SfiI (ggccattatggc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGCCCATATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCCGCGGCGGCGGATG-3' (30)BN-3',
(where B = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."

ORIGIN
Alignment Scores:
Pred. NO.: 3.37e-74 Length: 585
Score: 778.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-690-825-34 (1-142) x BG776624 (1-585)

Qy 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 46 ATGGGTGCCCCGAGTTGGCCCTGCTGCTGAGCCCTTCTCAAGGACCCGATCTCT 105
Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 106 ACATTCAGAACTGGCCCTTCTTGGAGGCTGGCCCTGCACCCCGAGCGGATGGCCGAG 165
Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60
Db 166 GCTGGCTTCATCACTCCCTCCCATGTGAACAGCCAGACTTGGCCAGTGTCTTCTGTC 225
Qy 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80
Db 226 TTCAAGGAGCTGGAAGCTGGAGCTGGAGCCAGATGACGACCCCATAGAGGAACATAAAAGCAT 285
Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
Db 286 TCGTCCCGTTGCGCTTCTTCTGTCGAAAGACAGTTTGAAGAATTAACCTTGTGTAA 345
Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
Db 346 TTTTGAACCTGGACAGAGAAAGAGCCAGCAACAAAAATTCGAAAGGAAACCAACATAAG 405
Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaA 140
Db 406 AAGAAAGAAATTTGAGGAATTCGGAAGAAAGTGGCCGTCCTGATCGAGCAGCTGGCTGCC 465
Qy 141 MetAsp 142
Db 466 ATGGAT 471

RESULT 12
AL603004 602663774F1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4808802 5',
LOCUS mRNA sequence.
DEFINITION DKFZp686k1620_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
AL603004
ACCESSION AL603004
VERSION AL603004.1 GI:15166510
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 602)  
 Wambutt, R., Heubner, D., Mewes, W., Weil, B. and Wiemann, S.  
 EST (Wambutt, R., Heubner, D., Mewes, W., Weil, B. and Wiemann, S.)  
 Unpublished (1999)  
 Contact: MIPS  
 MIPS  
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
 sequenced by AGORA (Berlin/Germany) within the cDNA sequencing  
 consortium of the German Genome Project.  
 No si sequence available.  
 This clone (DKFZ6686K1620) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
 Location/Qualifiers  
 1..602  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DKFZ6686K1620"  
 /dev\_stage="adult"  
 /lab\_host="DH108"  
 /clone\_lib="686 (synonym: hicc3)"  
 /note="Vector: pTriplEx2; Site\_1: SfiIA; Site\_2: SfiIB;  
 cDNA-collection"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 3 486-74 Length: 602  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-690-825-34 (1-142) x AL503004 (1-602)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 DB 45 ATGGTGGCCCGACGTTGGCCCTGGCTGGCAGCCCTTTCTCAAGACACCGCATCTCT 104  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 105 ACATTCAAGAACTGGCCCTTTCTGGAGGCTGGCCCTGCACCCCGAGCGATGGCCGAG 164  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
 DB 165 GCTGGCTTCATCCACTGCCCTCAGAGACGAGCCAGACTTGGCCCGAGTGTCTCTGTC 224  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80  
 DB 225 TTCAGGAGCTGGAGGCTGGAGCCAGATGACGACCCCATAGAGAAATATAAAGCAT 284  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 DB 285 TCCTCGGTGGCTTCTCTCTCTGTCACAGACGAGCTTTGAGAAATTAACCTTGGTGAA 344  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 DB 345 TTTTTCGAACTGGACAGAGAGAGCCAGAACAAAAATTGCAAGGAAACCAACATAG 404  
 QY 121 LysLysGluPheGluGluThrAlaLysValArgAlaIleGluGlnLeuAla 140  
 DB 405 AAGAAGAATTGAGGAACTGGAGAAAGTGGCGCGTCCCATCGAGCAGCTGGCTGCC 464  
 QY 141 MetAsp 142  
 DB 465 ATGAT 470

RESULT 13

BM841904 605 bp mRNA linear EST 06-MAR-2002  
 LOCUS K-EST0119223 S13KMS5 Homo sapiens cDNA clone S13KMS5-34-A12 5',  
 DEFINITION mRNA sequence.  
 BM841904  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 605)  
 Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
 Kim, Y.S.  
 21C Frontier Korean EST Project 2001  
 Unpublished (2002)  
 Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 34 row: A column: 12  
 High quality sequence stop: 605.  
 Location/Qualifiers  
 1..605  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S13KMS5-34-A12"  
 /tissue\_type="myeloma"  
 /cell\_line="KMS-5"  
 /lab\_host="Top10F"  
 /clone\_lib="S13KMS5"  
 /note="Vector: pCNS; Site\_1: EcoRI; Site\_2: NotI; The poly  
 (A) + RNA was dephosphorylated with bacterial alkaline  
 phosphatase (BAP) and then decapped with tobacco acid  
 pyrophosphatase (TAP). The decapped intact mRNA was  
 ligated with DNA-RNA linker including EcoR I site by  
 treatment of T4 RNA ligase and the first strand cDNA was  
 synthesized from oligo dt-selected mRNA by priming with  
 dt-tailed vector. The dt-tailed vector was adjusted to  
 have about 60nt. The cDNA vector was circularized with E.  
 coli DNA ligase after digestion of EcoRI which site is  
 also included in vector. An RNA strand converted to a DNA  
 strand by Okayama-Berg method. The obtained cDNA vectors  
 were used for transformation of competent cells E. coli  
 Top10F by electroporation method. The cDNA libraries  
 constructed by this method are full-length enriched cDNA  
 library."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 3 58-74 Length: 605  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-690-825-34 (1-142) x BM841904 (1-605)  
 QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 DB 78 ATGGTGGCCCGACGTTGGCCCTGGCTGGCAGCCCTTTCTCAAGACACCGCATCTCT 137  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 138 ACATTCAAGAACTGGCCCTTTCTGGAGGCTGGCCCTGCACCCCGAGCGATGGCCGAG 197  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60

Db	198	GCTGGCTTCATCCACTGCCCACTGGAACGAGCAGACTGGCCCACTGTTCTTCTGC	257
Qy	61	PhelysGluLeuGluGluYTrpGluProAspAspProIleGluGluHisLysLysHis	80
Db	258	TTCAAGGAGCTGGAGGCTGGAGCCAGATGACGACCCCATAGAGGAAACATAAAGCAT	317
Qy	81	SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu	100
Db	318	TCGTCGGGTTCGGCTTTCTTCTGTCAAGAAGCAGTTTCAAGAAATTAACCCCTGGTGA	377
Qy	101	PheLeuLysLeuAspArgGluAatGAlaLysAsnLysIleAlaLysGluThrAsnAsnLys	120
Db	378	TTTTTGAACCTGGACAGAGAAGAGCCAGAACAAAATTGCAAGGAAACCAACATPAG	437
Qy	121	LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla	140
Db	438	AAGAAAGAAATTTAGGAAATCTGCAAGAAAGTGGCGCTGCCATCGAGCAGCTGGCTGCC	497
Qy	141	MetAsp	142
Db	498	ATGGAT	503
RESULT 14			
LOCUS	AW409816		
DEFINITION	f102e10.x1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2961115 5', mRNA linear EST 29-JUN-2000		
ACCESSION	AW409816		
VERSION	AW409816.1 GI:6935357		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 613)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-r@mail.nih.gov		
	Tissue Procurement: AICC		
	cDNA Library Preparation: Ling Hong/Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: National Institutes of Health Intramural		
	Sequencing Center (NISC)		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	www.bio.llnl.gov/bbrp/image/image.html		
	Plate: LLCM55 row: I column: 20		
	Seq primer: -21m13 forward primer (ABI).		
FEATURES	Location/Qualifiers		
source	1..613		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:2961115"		
	/tissue_types="rhabdomyosarcoma"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH_MGC_17"		
	/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."		
ORIGIN			
Alignment Scores:			
Pred. No.:	3..56e-74	Length:	613
Score:	778.00	Matches:	142



2007 06:00:00 /T 5m ant

**00000000**

Search completed: August 16, 2004, 00:12:56  
Job time : 2792 secs

*This Page Blank (uspio)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 15, 2004, 16:29:32 ; Search time 457 Seconds  
(without alignments)  
1320.009 Million cell updates/sec

Title: US-09-690-825-34  
Perfect score: 778  
Sequence: 1 MGATLPPAWQFLKDHRS.....EFETAKVKVRAIEQLAAMD 142

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -MODEL=frame\_p2n.model -DEV=xlh  
-Q/cgn2\_1/USFTO\_spool/US09690825/runat\_11082004\_140951\_13523/app\_query.fasta\_1.327  
-DB=N\_Geneseq\_29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPECL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09690825 @CGN 1.1 470 @runat\_11082004\_140951\_13523 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT\_DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREATS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_29Jan04:.\*  
1: Geneseqn1980s:.\*  
2: Geneseqn1990s:.\*  
3: Geneseqn2000s:.\*  
4: Geneseqn2001as:.\*  
5: Geneseqn2001bs:.\*  
6: Geneseqn2002s:.\*  
7: Geneseqn2003as:.\*  
8: Geneseqn2003bs:.\*  
9: Geneseqn2003cs:.\*  
10: Geneseqn2004as:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	778	100.0	426	AAZ31037	AAZ31037 Open read
2	778	100.0	426	AAZ42859	AAZ42859 Survivin
3	778	100.0	429	ACF58071	ACF58071 Human sur
4	778	100.0	1024	ABZ83360	ABZ83360 Toxicolog
5	778	100.0	1619	AAA08903	AAA08903 Human sur
6	778	100.0	1619	AAZ21617	AAZ21617 DNA encod
7	778	100.0	1619	ABX76199	ABX76199 Lung canc
8	778	100.0	1619	AAZ53734	AAZ53734 Survivin

9	778	100.0	1619	7	AAZ49666	Human sur
10	778	100.0	1619	9	ADZ58228	Farnesyl
11	778	100.0	1662	7	ABZ58106	Human Sur
12	778	100.0	2404	7	ACA03988	CDNA down
C 13	748	96.1	1165	2	AAQ93052	Human EPR
14	721	92.7	740	7	ACA56320	Human sig
15	660	84.8	500	2	AAT72714	Mouse inh
16	660	84.8	955	5	AAZ21530	DNA encod
17	642	82.5	393	3	AAZ02317	Human sec
18	465	59.6	300	9	ADD33414	Mouse mit
19	382	49.1	14796	2	AAV27941	Survivin
20	382	49.1	14796	4	AAH47531	Human Her
21	382	49.1	14796	5	AAZ21523	DNA encod
22	382	49.1	14796	6	ABL68655	Kidney ca
23	382	49.1	14796	6	ABL66326	Lung canc
24	382	49.1	14796	6	ABL56663	Lung canc
25	382	49.1	14796	6	ABN96924	Gene #342
26	246.5	31.7	598	4	ABL09161	Drosophil
27	246.5	31.7	624	9	ADC06776	Fruit fly
28	233	29.9	121	7	ABZ68313	Nucleotid
29	219	28.1	403	2	AAZ31036	5' flanki
30	214	27.5	1100	4	AAF82458	Human sur
C 31	201.5	25.9	1050	9	ADB67958	Human lun
C 32	184.5	23.7	2736	4	ABL09180	Drosophil
C 33	184.5	23.7	2926	4	ABL09178	Drosophil
C 34	184.5	23.7	5749	4	ABL18270	Drosophil
C 35	160	20.6	14490	4	AAH20191	Human apo
36	158	20.3	14843	4	ABL10865	Drosophil
C 37	158	20.3	20672	4	ABL10864	Drosophil
38	157	20.2	1075	2	AAZ72715	C. elegan
39	155	19.9	5366	2	AAZ56273	Human apo
40	155	19.9	5366	2	AAZ58001	Gonadotro
41	155	19.9	5502	2	AAZ30092	Neuronal
42	155	19.9	5502	2	AAZ71263	Neuronal
43	155	19.9	5984	2	AAZ56272	Human apo
44	155	19.9	5984	2	AAZ58000	Gonadotro
45	155	19.9	5984	3	AAA39808	Human NAI

ALIGNMENTS

RESULT 1  
AAZ31037  
ID AAZ31037 standard; DNA; 426 BP.

XX AC AAZ31037;  
XX AC AAZ31037;  
DT 17-JAN-2000 (first entry)  
XX Open reading frame of human Survivin gene.

Survivin; open reading frame; IAP; apoptosis inhibitor; tumour cell;  
effector cell protease receptor-1; EPR-1; embryonic tissue; tubulin;  
mitotic spindle; abnormal cell growth; tumourigenesis; ss.

OS Homo sapiens.  
FX Key Location/Qualifiers  
FT CDS 1..426  
FT /tag= a  
FT /product= "Survivin"

FN WO3950440-A2.

XX PD 07-OCT-1999.

XX PF 01-APR-1999; 99WO-US007205.

XX PR 01-APR-1998; 98US-0080288P.

XX PA (UYFA ) UNIV YALE.

XX PI Altieri DC;

XX WPI; 1999-591329/50.  
 DR P-PSDB; AAY49080.  
 XX  
 PT Identifying agents that modulate Survivin interactions.  
 XX  
 PS Example; Fig 8B; 56pp; English.  
 XX  
 CC This sequence is the open reading frame of the human survivin gene.  
 CC Survivin is a 142 amino acid protein of approximately 16.5kD. Survivin is  
 CC a member of the IAP family of apoptosis inhibitors, and the gene is  
 CC located on chromosome 17q25. The nucleic acid sequence of Survivin is  
 CC related to the Effector Cell Protease Receptor-1 (EPR-1). Survivin is  
 CC expressed in tumour cells and embryonic tissue. The sequence is used in  
 CC the invention which involves identifying an agent that modulates  
 CC interactions between Survivin and tubulin. Agents that modulate  
 CC interaction between Survivin and tubulin or mitotic spindles may be  
 CC administered to a mammal to modulate biological or pathological processes  
 CC mediated by Survivin, for example Survivin-mediated inhibition of  
 CC cellular apoptosis. Thus an inhibiting agent may be used to block  
 CC abnormal cell growth, for example during tumourigenesis. An agent that  
 CC increases interaction may be used to extend cell growth in culture  
 XX  
 SQ Sequence 426 BP; 115 A; 112 C; 113 G; 86 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 9.17e-82 Length: 426  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-690-825-34 (1-142) x AAZ31037 (1-426)

QY 1 MetGlyAlaProThrLeuProToAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 DB 1 ATGGGTGCGCCGACGTTGGCCCTGCTGGCAGCCCTTCTCAAGGACCAACCCATCTCT 60  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 61 ACATTCAAGAACTGGCCCTTCTTGAGGGCTGGCCCTGCGCCCTGCGAGCGGATGGCCGAG 120  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 DB 121 GCTGGCTTCATCCACTGCCCTGACAGACGACGACCTTGGCCAGTGTCTTCTCTGC 180  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80  
 DB 181 TTCAAGAGCTGGAAGCTGGAGCCGACAGATGACGACCCCATAGAGGAACATAAAGCAT 240  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 DB 241 TCGTCCGGTTGGCTTTCCTTCTGTCAAGAACGAGTTGAAGAATTACCCCTTGTGGA 300  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 DB 301 TTTTGAACCTGGACAGAGAAGAGCCAGAACAAATTTGCAAGGAACCAACAATAAG 360  
 QY 121 LysLysGluPheGluGluThrAlaLysValArgAlaIleGluGlnLeuAla 140  
 DB 361 AAGAAAGATTGAGGAAATTCGGAAGAAAGTGGCCGCTGCCATCGAGCAGCTGGCTGCC 420  
 QY 141 MetAsp 142  
 DB 421 ATGGAT 426

RESULT 2

AAJ42859

ID AAL42859 standard; DNA; 426 BP.

XX

AC

XX

DT 05-AUG-2002 (first entry)  
 XX Survivin-like protein coding sequence 6.  
 XX  
 KW Survivin-like protein; diagnosis; screening; cancer; gene; ds;  
 KW apoptosis abnormality; gene therapy.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..426  
 FT /tag= a  
 FT /partial  
 FT /product= "survivin-like protein 6"  
 FT /note= "No stop codon is given"  
 XX  
 PN WO200233071-A1.  
 XX  
 PD 25-APR-2002.  
 XX  
 PF 16-OCT-2001; 2001WO-JP009071.  
 XX  
 PR 17-OCT-2000; 2000JP-00316721.  
 PR 20-DEC-2000; 2000JP-00386809.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Tanaka H, Kaleda I;  
 XX  
 DR WPI; 2002-435536/46.  
 DR P-PSDB; AAO14948.  
 XX  
 PT Baculovirus IAP repeat domain or RING-finger domain-containing survivin-  
 PT like polypeptides and encoded DNAs, applicable in diagnosis and screening  
 PT compounds for treating various cancers and apoptosis abnormality.  
 XX  
 PS Disclosure; Page 123-124; 136pp; Japanese.  
 XX  
 CC The invention comprises the amino acid and coding sequences of survivin-  
 CC like proteins. The survivin-like DNA and protein sequences are useful in  
 CC diagnostics and screening compounds for treating various cancers and  
 CC apoptosis abnormality, including gene therapy. The present DNA sequence  
 CC encodes a survivin-like protein of the invention  
 XX  
 SQ Sequence 426 BP; 115 A; 112 C; 113 G; 86 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 9.17e-82 Length: 426  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x AAL42859 (1-426)

QY 1 MetGlyAlaProThrLeuProToAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 DB 1 ATGGGTGCGCCGACGTTGGCCCTGCTGGCAGCCCTTCTCAAGGACCAACCCATCTCT 60  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 61 ACATTCAAGAACTGGCCCTTCTTGAGGGCTGGCCCTGCGCCCTGCGAGCGGATGGCCGAG 120  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 DB 121 GCTGGCTTCATCCACTGCCCTGACAGACGACGACCTTGGCCAGTGTCTTCTCTGC 180  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80  
 DB 181 TTCAAGAGCTGGAAGCTGGAGCCGACAGATGACGACCCCATAGAGGAACATAAAGCAT 240  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluGluLeuThrLeuGlyGlu 100

Db 241 TCGTCGGTGGCTTCCTTCCTGTCAGAGCAGTTTGAAGATTAAACCTTGGTGA 300  
Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLeuAlaLysGluThrAsnLys 120  
Db 301 TTTTGAACCTGGACAGAGAGAGCCCAAGAACAAAATTGCAAGAGGAAACCAACAATAAG 360  
Qy 121 LysLysGluPheGluThrAlaLysLysValArgArgAlaLysGluGlnLeuAlaAala 140  
Db 361 AAGAAGAATTGAGGAAACTGCGNAGAAAGTGGCCGTGCAATCGAGCAGCTGGCTGCC 420  
Qy 141 MetAsp 142  
Db 421 ATGGAT 426  
RESULT 3  
ID ACF58071 standard; DNA; 429 BP.  
AC ACF58071;  
XX  
DT 15-JAN-2004 (first entry)  
XX Human survivin protein encoding DNA.  
DE RasGAP; Aurora kinase; cytostatic; anti-tumoural; human; survivin; gene;  
KW ds.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..429  
FT FT /\*tag= a  
FT FT /product= "survivin"  
XX  
PN WO2003087395-A2.  
XX  
PD 23-OCT-2003.  
XX  
PF 15-APR-2003; 2003WO-IB002972.  
XX  
PR 15-APR-2002; 2002US-0372483P.  
XX  
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
XX  
PI Garbay C, Gigoux V, Canonis J, L'hoste S, Samson J;  
XX  
DR WPI: 2003-845340/78.  
DR P-PSDB; ABR61582.  
XX  
PT Identifying anti-tumoral compounds, comprises determining the capacity of  
PT a compound to inhibit interaction between RasGAP and Drosophila  
PT melanogaster (Dm) aurora kinase, or between RasGAP, Dm aurora kinase and  
PT aurora binding protein.  
XX  
PS Disclosure; Page 67; Opp; English.  
XX  
CC The invention relates to identifying a biologically active compound with  
CC anti-tumoural properties, where the compound is studied for its capacity  
CC to inhibit the interaction between: (a) RasGAP and the Drosophila  
CC melanogaster Aurora kinase or an orthologue of the kinase, or its  
CC fragment able to interact with RasGAP; or (b) RasGAP, an Aurora-binding  
CC protein and the D. melanogaster Aurora kinase or an orthologue of the  
CC kinase, or its fragment. The method is useful for identifying anti-  
CC tumoural agents. The present sequence represents a DNA encoding a human  
CC survivin protein, an aurora-binding protein  
XX  
SQ Sequence 429 BP; 116 A; 112 C; 114 G; 87 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 9.25e-82 Length: 429  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
US-09-690-825-34 (1-142) x ACF58071 (1-429)  
Qy 1 MetGlyAlaProThrLeuProProAlaThrGlnProPheLeuLysAspHisArgIleSer 20  
Db 1 ATGGTGCCCGGACGTTGCCCTTCCTGCTGCGAGCCCTTCTCAAGGACCACCGCATCTCT 60  
Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 61 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGGCTCGACCCCGGCGGATGGCCGAG 120  
Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
Db 121 GCTGGCTTCATCCACTGCCCTCACTGAGAACGAGCCAGACTTGGCCAGTGTTCCTTCTGC 180  
Qy 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluHisLysLysHis 80  
Db 181 TTCAGAGAGCTGGAAGGCTGGAGCCAGATGACCCCATAGAGGAACATAAAGCAT 240  
Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100  
Db 241 TCGTCCGTTGGCTTTCCTTCTGTCAAGAGCAGTTTGAAGATTAAACCTTGGTGA 300  
Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLeuAlaLysGluThrAsnLys 120  
Db 301 TTTTGAACCTGGACAGAGAGAGCCCAAGAACAAAATTGCAAGAGGAAACCAACAATAAG 360  
Qy 121 LysLysGluPheGluThrAlaLysLysValArgArgAlaLysGluGlnLeuAlaAala 140  
Db 361 AAGAAGAATTGAGGAAACTGCGNAGAAAGTGGCCGTGCAATCGAGCAGCTGGCTGCC 420  
Qy 141 MetAsp 142  
Db 421 ATGGAT 426  
RESULT 4  
ABZ83360  
ID ABZ83360 standard; cDNA; 1024 BP.  
XX  
AC ABZ83360;  
XX  
DT 14-MAY-2003 (first entry)  
XX  
DE Toxicologically relevant human nucleotide sequence #519.  
XX  
KW Toxicologically relevant gene; toxicological response; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2003016500-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 16-AUG-2002; 2002WO-US026514.  
XX  
PR 16-AUG-2001; 2001US-0313080P.  
XX  
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.  
XX  
PI Neft RE, Dunn RT, Adkins K, Pickett GG, Kier LD, Schweiser K;  
XX Alen P;  
XX  
DR WPI: 2003-268322/26.  
XX  
PT Determining a toxicological response to an agent, useful for screening of  
PT drugs, comprises comparing the expression profile of one or more human  
PT toxic response genes to a reference gene expression profile indicative of  
PT toxicity.  
XX  
PS Claim 1; Page 183; 455pp; English..



CC The invention relates to antisense oligonucleotides targeted to a nucleic  
CC acid molecule encoding human Survivin, where the antisense  
CC oligonucleotide inhibits the expression of human Survivin. These  
CC antisense oligonucleotides are used in the treatment of an animal  
CC suffering from a disease or condition associated with Survivin, e.g. a  
CC hyperproliferative condition such as cancer, and comprises administering  
CC a therapeutically or prophylactically effective amount of the antisense  
CC oligonucleotide so that expression of Survivin is inhibited. The  
CC oligonucleotide can also be used to treat a human suffering from a  
CC disease or condition characterised by a reduction in apoptosis comprising  
CC administering the antisense oligonucleotide to a human. In addition, the

```

PF 18-APR-2002; 2002WO-US012476.
XX
PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Aziz N, Murray R;
XX
DR WPI; 2003-093161/08.
DR P-PSDB; ABUS6477.
XX
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
PS Claim 22; Page 242; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridizes
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC invention
XX
SQ Sequence 1619 BP; 349 A; 392 C; 449 G; 429 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,49e-81 Length: 1619
Score: 778.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-690-825-34 (1-142) x ABX76199 (1-1619)

Qy 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 50 ATGGGTGCCCGACGTTGGCCCTGCTGCGACGCCCTTTCTCAAGGACCACCGCATCTCT 109
Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 110 ACATTCAGAACTGGCCCTCTTGGAGGGCTGGCGCTGCACCCCGAGCGGATGGCCGAG 169
Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db 170 GCTGGCTTCATCCACTGCCCACTGAGACGACGCGACACTTGGCCCGAGTGTCTTCTCTCG 229
Qy 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80
Db 230 TTCAAGGAGCTGGAAGCTGGGAGCCAGATGACGACCCCATAGAGGAACATAAAAGCAT 289
Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100

```

290 TCGTCGGTTCGGCTTTCTCTCTCAAGAGCAGTTTGAAGAATTAAACCTTGGTGA 349

101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120

350 TTTTGAACCTGGACAGAGAAAGGCCCAAGACAAATTGCAAGGAAACCAACAATAAG 409

121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140

410 AAGAAAGAAATTGAGGAACTCGGAAGAAAGTGGCGCGTGCATCGAGCAGCTGGCTGCC 469

141 MetAsp 142

470 ATGGAT 475

RESULT 8

AAL53734

ID AAL53734 standard; RNA; 1619 BP.

XX

AC AAL53734;

XX

DT 07-FEB-2003 (first entry)

XX

DE Survivin DNA sequence SEQ ID No 28.

XX

KW Target RNA; target RNA-support-attached test compound; flow cytometry;

KW mass spectrometry; high-throughput screening; ds.

XX

OS Homo sapiens.

XX

PN WO200283837-A1.

XX

PD 24-OCT-2002.

XX

PF 11-APR-2002; 2002WO-US011758.

XX

PR 11-APR-2001; 2001US-0282966P.

XX

PA (PTCT-) PTC THERAPEUTICS INC.

XX

PI Almstead NG;

XX

DR WPI; 2003-075534/07.

XX

PT Identifying a test compound that binds to a target RNA molecule by

PT separating the detectably labeled target RNA-support-attached test

PT compound complex from uncomplexed target RNA molecules and test compounds

PT by flow cytometry.

XX

Example; Page 81; 131pp; English.

XX

The invention relates to a novel method for identifying a test compound

CC that binds to a target RNA molecule comprising separating the detectably

CC labeled target RNA-support-attached test compound complex from

CC uncomplexed target RNA molecules and test compounds. The separating

CC process is carried out by flow cytometry and determining a structure of

CC the type of test compound of the RNA-support-attached test compound

CC complex by mass spectrometry. The method is useful for high-throughput

CC screening of libraries of compounds to identify pharmaceutical leads.

CC This polynucleotide sequence represents a DNA sequence related to the

CC detecting method of the invention

XX

SQ Sequence 1619 BP; 349 A; 392 C; 449 G; 429 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,49e-81 Length: 1619

Score: 778.00 Matches: 142

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 7 Gaps: 0

US-09-690-825-34 (1-142) x AAL53734 (1-1619)

81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100



QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 50 ATGGGTGCCCCACGTTGCCCTCCCTGGCAGCCCTTCTCAAGGACCCACCGCATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 110 ACATTCAAGAAGTGGCCCTCTTTGAGGGCTGCGCTGCACCCGAGCGGATGCGCGAG 169  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 170 GTGTGGCTTCATCCACTGCCCCACTGAGAACGAGCCAGACTTGGCCCACTGTTCTTCTG 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
DB 230 TTCAGGAGCTGGAAGCTGGAGCGCAGTACGACCCCTAGAGAACATATAAAGCAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 290 TCGTCGGTGGCGCTTCTCTCTGCAAGAAGCAGTTTGAAGAATTAAACCTTGGTGAA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLysLysLysLysLys 120  
DB 350 TTTTGTAAACTGGACAGAGAAAGCCAAAGAAATTCGAAGGAAACCAACATTAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
DB 410 AAGAAAGAAATTTGAGGAACCTGCGAAGAAAGTGGCGCTGCCATCGAGCAGCTGCTGCC 469  
QY 141 MetAsp 142  
DB 470 ATGGAT 475  
RESULT 9  
AAD49666  
ID AAD49666 standard; DNA; 1619 BP.  
XX AC AAD49666;  
XX DT 24-MAR-2003 (first entry)  
XX DE Human survivin DNA.  
XX KW Amyloidosis; haemophilia; Alzheimer's disease; atherosclerosis; cancer;  
KW gigantism; dwarfism; hypothyroidism; hyperthyroidism; cystic fibrosis;  
KW autoimmune disorder; aging; inflammation; diabetes; obesity; anorectic;  
KW neurodegenerative disorder; Parkinson's disease; gene therapy; viricide;  
KW haemostatic; antibacterial; neutropic; neuroprotective; cytostatic; XIAP;  
KW fungicide; human; survivin; gene; ds.  
OS Homo sapiens.  
XX WO200283953-A1.  
XX PN 24-OCT-2002.  
XX PD 11-APR-2002; 2002WO-US011757.  
XX PF 11-APR-2001; 2001US-0282965P.  
XX PR (PTCT-) PTC THERAPEUTICS INC.  
XX PA Rando R, Welch E;  
XX PI WPI; 2003-075561/07.  
XX DR Identifying a test compound that binds to a target RNA molecule for  
XX PT treating or preventing amyloidosis, hemophilia, cancer, gigantism,  
XX PT diabetes, by contacting a detectably labeled target RNA molecule with a  
XX PT library of test compounds.  
XX Example; Page 89-90; 152pp; English.  
XX PS The invention relates to a method for identifying a test compound that

CC binds to a target RNA molecule, which comprises contacting a detectably  
CC labelled target RNA molecule with a library of test compounds under  
CC conditions that permit direct binding of the labelled target RNA to a  
CC member of the library of test compounds so that a detectably labeled  
CC target RNA: test compound complex is formed. The method is useful for  
CC screening libraries of compounds for those that are selectively bind to a  
CC pre-selected target RNA. The compounds are useful for inhibiting the  
CC formation of a specific bound RNA: host cell factor complexes in vivo.  
CC They are also useful for treating or preventing diseases associated with  
CC overproduction or decreased protein function, such as amyloidosis;  
CC haemophilia, Alzheimer's disease, atherosclerosis, cancer, gigantism,  
CC dwarfism, hypothyroidism, hyperthyroidism, autoimmune disorders, aging,  
CC inflammation, cystic fibrosis, diabetes, obesity, neurodegenerative  
CC disorders, Parkinson's disease or infections (bacterial, viral, fungal).  
CC The invention is also used in gene therapy. The present sequence is human  
CC survivin DNA. This sequence is used to illustrate the method of the  
CC invention  
XX  
SQ Sequence 1619 BP; 349 A; 392 C; 449 G; 429 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 5,49e-81 Length: 1619  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-09-630-825-34 (1-142) x AAD49666 (1-1619)  
QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 50 ATGGGTGCCCCACGTTGCCCTCCCTGGCAGCCCTTCTCAAGGACCCACCGCATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 110 ACATTCAAGAAGTGGCCCTCTTTGAGGGCTGCGCTGCACCCGAGCGGATGCGCGAG 169  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 170 GTCGCTTCATCCACTGCCCCACTGAGAACGAGCCAGACTTGGCCCACTGTTCTTCTG 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
DB 230 TTCAGGAGCTGGAAGCTGGAGCGCAGTACGACCCCTAGAGAACATATAAAGCAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 290 TCGTCGGTGGCGCTTCTCTCTGCAAGAAGCAGTTTGAAGAATTAAACCTTGGTGAA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLysLysLysLysLys 120  
DB 350 TTTTGTAAACTGGACAGAGAAAGCCAAAGAAATTCGAAGGAAACCAACATTAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
DB 410 AAGAAAGAAATTTGAGGAACCTGCGAAGAAAGTGGCGCTGCCATCGAGCAGCTGCTGCC 469  
QY 141 MetAsp 142  
DB 470 ATGGAT 475  
RESULT 10  
ADE85228  
ID ADE85228 standard; DNA; 1619 BP.  
XX AC ADE85228;  
XX DT 29-JAN-2004 (first entry)  
XX DE Farnesyl transferase inhibitor modulated leukemia associated gene #447.  
XX KW ss; cytostatic; farnesyl transferase inhibitor; gene expression;

KW quinolinone; leukemia; cancer.

XX Homo sapiens.

XX WO2003038129-A2.

XX 08-MAY-2003.

XX 30-OCT-2002; 2002WO-US034784.

XX 30-OCT-2001; 2001US-0338997P.

XX 30-OCT-2001; 2001US-0340081P.

XX 30-OCT-2001; 2001US-0340938P.

XX 30-OCT-2001; 2001US-0341012P.

XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.

XX Raponi M;

XX WPI; 2003-513497/48.

XX Determining whether a patient will respond to treatment with a farnesyl transferase inhibitor, by analyzing the expression of gene that is differentially modulated in the presence of the inhibitor.

XX Disclosure; SEQ ID NO 447; 346pp; English.

XX The invention relates to a method of determining whether a patient will respond to treatment with a farnesyl transferase inhibitor (FTI), by analyzing the expression of gene that is differentially modulated in the presence of an FTI. The method is useful for determining whether a patient will respond to treatment with a FTI such as (B)-6-[amino(4-chlorophenyl)(1-methyl-1H-imidazol-5-yl)methyl]-4-(3-chlorophenyl)-1-methyl-2-(1H)quinolinone, monitoring the therapy of a patient, treating a patient with leukemia with FTI if the analysis indicates that the patient will respond. This sequence corresponds to a gene whose expression may be modulated in the presence of FTI.

XX Sequence 1619 BP; 349 A; 392 C; 449 G; 429 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 5,49e-81 Length: 1619  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x ADE85228 (1-1619)

QY 1 MetGlyAlaProThrLeuProAlaTrpClnProPheLeuLysAspHisAlaGleSer 20  
DB 50 ATGGTGCCCGCAGCTGGCCCTTCTCTGCGAGCCCTTTCTCAAGGACCAACCCGATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 110 ACATTCAAGAACTGGCCCTTCTTGAGGCGTGGCCCTGCACCCCGAGCGATGGCGGAG 169  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 170 GCTGCTTCATCCACTGCCCTCCCTGAGACGAGCCAGCTGGCCGAGTTCCTCTGTCG 229  
QY 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluHisLysLysHis 80  
DB 230 TTCAAGGAGCTGGAAGCTGGAGCCAGATGACGACCCCATAGAGGAACATAAAGCAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluLeuLeuThrLeuGlyGlu 100  
DB 290 TCGTCCCGTTCGGCTTTCCTTCTGTCGAAGAGCAGTTTGAAGAATTAACCCCTGGTGAA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
DB 350 TTTTGAACCTGGACAGAGAAGAGCCAGAACAAATTCGAAGGAACCAACAATAG 409

QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140  
DB 410 AAGAAGAAATTGAGAACTGCGAAGAAAGTGGCCGTGCGATCGAGCAGCTGGCTGCC 469  
QY 141 MetAsp 142  
DB 470 ATGGAT 475

RESULT 11

ABZ58106

ID ABZ58106 standard; cDNA; 1662 BP.

XX AC ABZ58106;

XX DT 22-APR-2003 (first entry)

XX Human Survivin cDNA.

XX Human; Survivin; inhibitor of apoptosis; cell death; apoptosis; cancer;

XX cytostatic; cardiant; neuroprotective; gene therapy; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 43..471

XX FT /tag= a

XX FT /product= "Survivin"

XX WO2003004606-A2.

XX 16-JAN-2003.

XX 03-JUL-2002; 2002WO-US021002.

XX 03-JUL-2001; 2001US-00898158.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Troy CM, Shelanski ML;

XX WPI; 2003-210351/20.

XX P-PSDB; ABP72162.

XX New nucleic acid encoding an inhibitor-of-apoptosis protein, useful for

XX treating cancer, neurodegenerative disorder or cardiomyopathy.

XX Disclosure; Fig 20B; 124pp; English.

XX The present sequence is that of cDNA encoding human Survivin protein. The

XX invention provides a nucleic acid, such as an antisense oligonucleotide,

XX which specifically hybridizes to a nucleic acid encoding an inhibitor of

XX apoptosis protein. Survivin is an example of an inhibitor of apoptosis or

XX inducer of cell death protein of the invention. A claimed method for

XX inducing a cell's death comprises contacting the cell with the nucleic

XX acid under conditions permitting the nucleic acid to enter the cell,

XX especially the use of a vector, liposome, or a mechanical or electrical

XX means. The method is used to treat acute lymphocytic leukaemia, acute

XX myelogenous leukaemia, lung cancer, breast cancer, ovarian cancer,

XX prostate cancer, lymphoma, Hodgkin's disease, malignant melanoma,

XX neuroblastoma, renal cell carcinoma and squamous cell carcinoma (all

XX claimed). The invention also provides a second nucleic acid, which

XX specifically hybridizes to a nucleic acid encoding a protein, other than

XX caspase-2, that induces cell death. A claimed method for inhibiting a

XX cell's death comprises contacting the cell with the nucleic acid under

XX conditions permitting the nucleic acid to enter the cell. The method is

XX used to treat a neurodegenerative disorder (especially a brain disorder

XX or central nervous system disorder), or a heart disorder (especially

XX cardiomyopathy) in a human (all claimed)

XX Sequence 1662 BP; 396 A; 389 C; 451 G; 426 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 5,698-81 Length: 1662  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 7 Gaps: 0

US-09-690-825-34 (1-142) x ABZ59106 (1-1662)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 43 ATGGGTGCGCGAGCTTGCCTGCGAGCGCTTTCTCAAGGACCCACCGATCTCT 102  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 103 ACATTCAGAACTGCGCTTTCTTGAGGGCTGCGCTGCACCCCGAGCGATGGCGAG 162  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
DB 163 GCTGCTTCATCCACTGCGCCACTGAGAACGAGCGAGACTTTGGCCAGTGTTCCTCTGC 222  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProlleGluGluHisLysLysHis 80  
DB 223 TTCAAGGAGCTGGAAGGTGGAGCGCAGATCAGACCCCATAGAGGAACATAAAAGCAT 282  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 283 TCGTCGGTTCGCTTCTCTGCAAGAGCAGTTTGAGAAATTAACCTTTGGTGAA 342  
QY 101 PheLysLysLeuAspArgGluArgAlaLysLysLysLysLysLysLysLysLys 120  
DB 343 TTTTGTAACTGGACAG 402  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaLysLysLysLysLysLys 140  
DB 403 AAGAAAGAAATTTGAGAAACTCGGAAGAAAGTGGCCCGTCCATCGACAGCTGGCTGCC 462  
QY 141 MetAsp 142  
DB 463 ATGGAT 468

RESULT 12  
ACA03988  
ID ACA03988 standard; cDNA; 2404 BP.

XX  
AC ACA03988;  
XX  
DT 27-MAY-2003 (first entry)  
XX  
DE cDNA downregulated in senescent cells Incyte ID NO: 251651.4.

KW Human; senescence; s6; gene; cancer; proliferative disorder; leukaemia;  
KW adenocarcinoma; lymphoma; melanoma; myeloma; sarcoma; teratocarcinoma;  
KW adrenal gland cancer; bladder cancer; bone cancer; bone marrow cancer;  
KW brain cancer; breast cancer; cervical cancer; colon cancer; heart cancer;  
KW oesophageal cancer; gall bladder cancer; ganglial cancer; kidney cancer;  
KW liver cancer; lung cancer; muscle cancer; ovarian cancer; penile cancer;  
KW pancreatic cancer; parathyroid gland cancer; prostate cancer;  
KW salivary gland cancer; skin cancer; small intestine cancer;  
KW spleen cancer; stomach cancer; testicular cancer; thymic cancer;  
KW thyroid cancer; uterine cancer.

OS Homo sapiens.  
XX  
XX US2002192679-A1.  
XX  
XX 19-DEC-2002.  
XX  
XX 07-FEB-2002; 2002US-00071766.  
XX  
XX 09-FEB-2001; 2001US-0268380P.  
XX  
XX (CHEN/) CHEN H.

XX Chen H;  
PI  
XX WPI: 2003-328858/31.  
DR  
XX  
PT New combination comprising cDNAs or their complements, useful for  
PT detecting changes in expression of genes encoding proteins associated  
PT with senescence, and in diagnosing, staging or treating proliferative  
PT diseases, e.g. cancer.  
XX  
PS Example 13; Page 135-136; 195pp; English.  
XX  
CC The invention relates to a combination comprising a plurality of cDNAs,  
CC or their complements that are differentially expressed in cancer and  
CC other proliferative disorders. The combination is useful in detecting  
CC changes in expression of genes encoding proteins that are associated with  
CC senescence and in diagnosing, staging, treating, or monitoring the  
CC progression or treatment of subjects with proliferative diseases such as  
CC cancer e.g. adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma,  
CC sarcoma, teratocarcinoma; cancer of the adrenal gland, bladder, bone,  
CC bone marrow, brain, breast, cervix, colon, oesophagus, gall bladder,  
CC ganglia, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid  
CC gland, penis, prostate, salivary glands, skin, small intestine, spleen,  
CC stomach, testis, thymus, thyroid and uterus. The present sequence  
CC represents cDNA of genes that are downregulated in senescent cells  
XX  
SQ Sequence 2404 BP; 525 A; 588 C; 634 G; 628 T; 0 U; 29 Other;

Alignment Scores:  
Pred. No.: 9,338-81 Length: 2404  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 7 Gaps: 0

US-09-690-825-34 (1-142) x ACA03988 (1-2404)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 56 ATGGGTGCGCGAGCTTGCCTGCGAGCGCTTTCTCAAGGACCCACCGATCTCT 115  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 116 ACATTCAGAACTGCGCTTTCTTGAGGGCTGCGCTGCACCCCGAGCGATGGCGAG 175  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
DB 176 GCTGCTTCATCCACTGCGCCACTGAGAACGAGCGAGACTTTGCCCGAGTGTTCCTCTGC 235  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProlleGluGluHisLysLysHis 80  
DB 236 TTCAAGGAGCTGGAAGGTGGAGCGCAGATGACGCCCATAGAGGAACATAAAAGCAT 295  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 296 TCGTCCGTTGCGCTTTCTTCTGTCAGAGCAGGTTTGAAGAAATTAACCTTTGGTGAA 355  
QY 101 PheLysLysLeuAspArgGluArgAlaLysAsnLysLysLysLysLysLysLys 120  
DB 356 TTTTGTAACTGGACAG 415  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaLysLysLysLysLysLys 140  
DB 416 AAGAAAGAAATTTGAGGAACCTGCGAGAAAGTGGCCCGTCCATCGAGCAGCTGGCTGCC 475  
QY 141 MetAsp 142  
DB 476 ATGGAT 481

RESULT 13  
AAQ93052/c  
ID AAQ93052 standard; cDNA; 1165 BP.

```

XX AAQ93052;
AC 25-MAR-2003 (revised)
DT 07-DEC-1995 (first entry)
XX Human EPR-1 cDNA.
DE Effector cell protease receptor-1; EPR-1; lymphocyte;
KW lymphoproliferative disorder; ss.
KW Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 37..1050
FT /tag= a
XX
XX WO9520655-A1.
XX
XX 03-AUG-1995.
XX
XX 18-JAN-1995; 95WO-US000666.
XX
XX 28-JAN-1994; 94US-00189309.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX Altieri DC;
XX
XX WPI; 1995-275440/36.
XX
XX P-PSDB; AAR76744.
XX
XX New effector cell protease receptor-1 protein - used to develop prods.
XX for inducing lymphocyte proliferation or diagnosis and treatment of
XX lymphoproliferative disorders.
XX
XX Claim 10; Page 154-155; 181pp; English.
XX
XX Immunoscreening of a human lymphocyte expression library in lambda- gt11
XX using a MAB raised against human EPR-1 yielded clone lambda-104. This was
XX used as a probe in further screening of human cDNA libraries, and a
XX consensus EPR-1 cDNA sequence was deduced from 28 isolated clones. The
XX cDNA can be used for prodn. of recombinant EPR-1. (Updated on 25-MAR-2003
XX to correct PN field.)
XX
XX Sequence 1165 BP; 293 A; 331 C; 300 G; 241 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1.15e-77 Length: 1165
Score: 748.00 Matches: 1165
Percent Similarity: 99.30% Conservative: 1
Best Local Similarity: 98.59% Mismatches: 1
Query Match: 96.14% Indels: 1
DB: 2 Gaps: 0
US-09-690-825-34 (1-142) x AAQ93052 (1-1165)
QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
DB 1114 ATGGGTGCCCGGAGGTGCCCCCTGCTGGCAGCCCTTTCTCAAGACACCACCGCATCTCT 1055
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGlnArgMetAlaGlu 40
DB 1054 ACATTCAAGAACTGGCCCTTTCTTGGAGGCTGGCG-TGCACCCCGAGCGGATGGCCGAG 996
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB 995 GCTGGCTTCATCCATGCCCATCTGAGACGACGACGATTTGGCCAGTGTCTTCTTCG 936
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80
DB 935 TTCAGGAGCTGGAGCTGGGAGCCAGATGACGACCCCATAGAGGAACATAAAAGCAT 876

```

81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
875 TCGTCCGGTTCGGCTTTCTCTGTCAGGAAGCAGTTTGAGGAATTAAACCTTGGTGAA 816  
101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
815 TTTTGAACCTGGACAGAGAAAGAGCCAAAGAAAATTGCAAGAGAAACCAACATAAG 756  
121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
755 AAGAAAGAAATTGAGAAACTGCGAAGAAAGTGGCGTGCATCGAGCAGCTGCCTCG 696  
141 MetAsp 142  
695 ATGGAT 690

RESULT 14  
ACA56320  
ID ACA56320 standard; cDNA; 740 BP.  
XX  
XX ACA56320;  
AC  
XX  
DT 06-JUN-2003 (first entry)  
XX  
DE Human signalling pathway polynucleotide probe SEQ ID NO 918.  
XX  
XX Human; probe; ss; array element; Parkinson's disease;  
KW signalling pathway population; cancer; adenocarcinoma; leukaemia;  
XX immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.  
XX  
XX Homo sapiens.  
XX  
XX US6500938-B1.  
XX  
XX 31-DEC-2002.  
XX  
XX 30-JAN-1998; 98US-00016434.  
XX  
XX 30-JAN-1998; 98US-00016434.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Au-Young J, Seilhamer JJ;  
XX  
XX WPI; 2003-352189/33.  
XX  
XX Combination of polynucleotide probes, useful as array elements in a  
PT microarray for monitoring the expression of a number of target  
PT polynucleotides.  
XX  
XX Claim 1; SEQ ID NO 918; 65pp; English.  
XX  
XX The invention relates to a combination which, comprises a number of  
XX polynucleotide probes comprising a sequence selected from one of the 1490  
XX sequences mentioned in the specification. The combination is useful as an  
XX array element in a microarray for monitoring the expression of a number  
XX of target polynucleotides. The microarray is particularly useful in the  
XX diagnosis and treatment of cancer and immunopathology and neuropathology.  
XX The microarray is useful in diagnostics and treatment regimens, drug  
XX discovery and development, toxicological and carcinogenicity studies,  
XX forensics and pharmacogenomics. The microarray is also useful for  
XX monitoring progression of diseases and for developing sophisticated  
XX profiles for the effects of currently available therapeutic drugs. The  
XX combination is also useful for purifying a subpopulation of mRNAs, cDNAs  
XX and genomic fragments and in research and diagnostic applications. The  
XX array can detect changes in expression in a large number of genes coding  
XX for different signalling pathway populations which can be used to diagnose  
XX various diseases including cancer e.g. adenocarcinoma and leukaemia;  
XX immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease  
XX and Parkinson's disease. The present sequence represents a polynucleotide  
XX probe of the invention. Note: The sequence data for this patent did not  
XX form part of the printed specification but was obtained in electronic  
XX format directly from USPTO at

```
CC seqdata.uspto.gov/sequence.html?DocID=06500938B1
XX SQ Sequence 740 BP; 170 A; 186 C; 189 G; 184 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.: 9,096-75 Length: 740
Score: 721.00 Matches: 135
Percent Similarity: 95.74% Conservative: 0
Best Local Similarity: 95.74% Mismatches: 6
Query Match: 92.67% Indels: 1
DB: 7 Gaps: 0

US-09-690-825-34 (1-142) x ACA56320 (1-740)
QY 2 GlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSerThr 21
DB 22 GGGTGGGNCNCCTTGGCACT-CCCTGGGNTCCCTTCTCAAGGACCACCGCATCTCTACA 80
QY 22 PheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGluAla 41
DB 81 TTCAAGAACTGGCCCTTCTTGGAGGGCTGCGCTGCACCCCGGAGCGATGCCGAGGCT 140
QY 42 GlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCysPhe 61
DB 141 GCGTTCATCCATGCGCCCACTGANAACGAGCGAGACTTGGCCAGTGTCTTCTGCTTC 200
QY 62 LysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHisSer 81
DB 201 AAGGAGCTGGAGGCTGGGAGCGAGATGACGACCCCATAGAGGAACATAAAAGCATTCG 260
QY 82 SerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGluPhe 101
DB 261 TCCGGTGGCGCTTCTTCTGTCAGAGAGCGAGTTTGAAGAAATTAAACCTTGGTGAATTT 320
QY 102 LeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLysLys 121
DB 321 TTGAAACTGGACAGAGAAAGAGCCAAAGAAATTTGCAAGGAAACCAACCAATTAAGAAG 380
QY 122 LysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaMet 141
DB 381 AAAGAATTTGAGGAACTCGGAAGAAAGTGCGCCGTCATCGAGAGCTGCTGCCCATG 440
QY 142 Asp 142
DB 441 GAT 443

RESULT 15
ID AAT72714
XX AC AAT72714;
XX AC AAT72714;
DT 16-SEP-1997 (first entry)
DE Mouse inhibitor of apoptosis protein homologue MIHD DNA.
DE DE
XX KW Inhibitor of apoptosis protein; IAP; mammalian IAP homologue; MIHD;
XX KW degenerative disease; infectious disease; autoimmune disease; cancer;
XX KW gene therapy; diagnosis; ss.
XX OS Mus musculus.
XX PH Key Location/Qualifiers
XX FT 48..470
XX FT /*tag= a
XX FN WO9723501-A1.
XX PD 03-JUL-1997.
XX PF 20-DEC-1996; 96WO-AU000827.
XX PR 22-DEC-1995; 95AU-00007275.

XX PA (AMRA-) AMRAD OPERATIONS PTY LTD.
XX PI Vaux DL;
XX WPI; 1997-350966/32.
DR P-PSDB; AAM19749.
XX Isolated protein homologues of viral inhibitors of apoptosis - used to
PT modulate apoptosis for treatment of degenerative, infectious or
PT autoimmune diseases and cancer.
XX Claim 28; Page 71-72; 136pp; English.
XX An isolated nucleic acid molecule (AAT72714) codes for mammalian IAP
CC homologue D (MIHD) (AAM19749), a murine homologue of baculovirus
CC inhibitor of apoptosis protein (IAP). It was isolated following a
CC database search for sequences showing homology to a consensus sequence
CC (see also AAM19744) for IAP homologues. Animal IAP homologue nucleic
CC acids (see also AAT72710-13 and AAT72715-17) can be used to produce
CC polypeptides useful in methods for modulating apoptosis in animal cells,
CC specifically for treatment, by inhibition, of degenerative and infectious
CC disease or, by promotion, of cancer and autoimmune disease, and can be
CC used for gene therapy of these diseases
XX SQ Sequence 500 BP; 134 A; 131 C; 135 G; 100 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7,486-68 Length: 500
Score: 660.00 Matches: 118
Percent Similarity: 92.86% Conservative: 12
Best Local Similarity: 84.29% Mismatches: 10
Query Match: 84.83% Indels: 0
DB: 2 Gaps: 0

US-09-690-825-34 (1-142) x AAT72714 (1-500)
QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
DB 48 ATGGGAGCTCGCGCGCTGCCCGCCAGATCTGGCAGCTGTACCTCAAGAACTACCGCATGCC 107
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 108 ACCTTCAAGAACTGGCCCTTCTGAGGAGCTGGCCCTGCACCCCGATAGAGGAGCATAGAAGCAC 167
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB 168 GCTGGCTTCATCACTGCCCTACCGAAGACGAGCTGATTTGGCCCGAGTGTCTTCTGTC 227
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80
DB 228 TTTAAGGAATTGGAGGCTGGAAACCCGATGACAAACCCGATAGAGGAGCATAGAAGCAC 287
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB 288 TCCCGCTGGCTGGCGCTTCTCCTCCTCAAGAGAGAGATGGAAGAACTAACCGTCAGTGAA 347
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
DB 348 TTTCTTGAACCTGGACACAGAGAGAGCCAGACAAATTCGAAAGGAGAGACCAACACAAG 407
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140
DB 408 CAAAAGAGATTGGAAGAGACTGCAAGAGACTACCCGCTCAGTCAATTGAGAGCTGGTGCC 467

Search completed: August 15, 2004, 22:24:53
Job time : 463 secs
```

This Page Blank (uspto)

```

DEFINITION Survivin-like polypeptide and its DNA.
ACCESSION BD185366
VERSION BD185366.1 GI:31877566
KEYWORDS JP 2002355062-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 426)
AUTHORS Tanaka,H. and Kaieda,I.
TITLE Survivin-like polypeptide and its DNA
JOURNAL Patent: JP 2002355062-A 6 10-DEC-2002;
TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
PN JP 2002355062-A/6
PD 10-DEC-2002
PF 16-OCT-2001 JP 2001318533
PI HIROSHI TANAKA,ISAO KAIEDA
PC C12N15/09,A61K31/7088,A61K38/55,A61K39/395,A61K39/395,A61K48/
PC 00,A61P35/00,
PC A61P43/00,A61P43/00,C07K14/82,C07K16/32,C12N1/15,C12N1/19,PC
C12N1/21,
PC C12N5/10,C12P21/02,C12Q1/02,C12Q1/68,G01N33/15,G01N33/50,PC
G01N33/53,
PC G01N33/53,G01N33/566,G01N33/574,C12N15/00,C12N5/00,A61K37/64
CC Survivin-like polypeptide and its DNA
FH Key Location/Qualifiers
FT source 1..426
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..426
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 3,01e-77 Length: 426
Score: 778.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-690-825-34 (1-142) x BD185366 (1-426)
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
DB 1 ATGGGTGCCCCGACGTTGCCCTGCTGGCAGCCCTTTCTCAAGGACCCACCGCATCTCT 60
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 61 ACATTCAAGAACTGGCCCTTCTTTGGAGGGCTGCGCTGCACCCCGGAGCGGATGCCCGAG 120
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60
DB 121 GCTGGCTTCTACCTGCCCCCTGAGACGACGACGAGCTTGGCCAGTGTTCTTCCTGCG 180
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80
DB 181 TTCAAGGAGCTGGAAGCTGGGAGCGAGATGACGACCCCATAGAGGAACATAAAAGCAT 240
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100
DB 241 TCGTCCGGTTGCGCTTCTCTCTGTCAGAACAGATTGAAGAAATTAACCCCTTGGTGAA 300
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120
DB 301 TTTTGAACCTGACAGAGAAAGAGCCAGAACAAAAATTGCAAGGAAACCAACAATAAG 360
QY 121 LysLysGluPheGluThrAlaLysLysValArgAlaIleGluGlnLeuAla 140
DB 361 AAGAAAGAAATTTGAGGAACTGCGAAGAAAGTGGCCGCTGCCATCGAGCAGCTGGCTGCC 420

```

```

QY 141 MetAsp 142
DB 421 ATGGAT 426
RESULT 4
AR099347
LOCUS AR099347 1619 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6077709.
ACCESSION AR099347
VERSION AR099347.1 GI:12809113
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1619)
AUTHORS Bennett,C.Frank., Ackermann,E.J., Swayze,E.E. and Cowsett,L.M.
TITLE Antisense modulation of Survivin expression
JOURNAL Patent: US 6077709-A 1 20-JUN-2000;
FEATURES
source
1..1619
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.38e-76 Length: 1619
Score: 778.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-690-825-34 (1-142) x AR099347 (1-1619)
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
DB 50 ATGGGTGCCCCGACGTTGCCCTGCTGGCAGCCCTTTCTCAAGGACCCACCGCATCTCT 109
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 110 ACATTCAAGAACTGGCCCTTCTTTGGAGGGCTGCGCTGCACCCCGGAGCGGATGCCCGAG 169
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60
DB 170 GCTGGCTTCTACCTGCCCCCTGAGACGACGACGAGCTTGGCCAGTGTTCTTCCTGCG 229
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80
DB 230 TTCAAGGAGCTGGAAGCTGGGAGCCAGATGACGACCCCATAGAGGAACATAAAAGCAT 289
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100
DB 290 TCGTCCGGTTGCGCTTCTCTCTGTCAGAACAGATTGAAGAAATTAACCCCTTGGTGAA 349
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120
DB 350 TTTTGAACCTGACAGAGAAAGAGCCAGAACAAAAATTGCAAGGAAACCAACAATAAG 409
QY 121 LysLysGluPheGluThrAlaLysLysValArgAlaIleGluGlnLeuAla 140
DB 410 AAGAAAGAAATTTGAGGAACTGCGAAGAAAGTGGCCGCTGCCATCGAGCAGCTGGCTGCC 469
QY 141 MetAsp 142
DB 470 ATGGAT 475
RESULT 5
BD273550
LOCUS BD273550 1619 bp DNA linear PAT 17-JUL-2003
DEFINITION Antisense modulation of survivin expression.
ACCESSION BD273550
VERSION BD273550.1 GI:33083318

```

**KEYWORDS** JP 2002539073-A/1.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
**REFERENCE** 1 (bases 1 to 1619)  
**AUTHORS** Bennett, F.C., Ackermann, E.J., Swayze, E.E. and Cowsett, L.M.  
**TITLE** Antisense modulation of survivin expression  
**JOURNAL** Patent: JP 2002539073-A 1 19-NOV-2002;  
**COMMENT** ISIS PHARMACEUTICALS INC  
**OS** Homo sapiens (human)  
**PN** JP 2002539073-A/1  
**PD** 19-NOV-2002  
**PF** 23-SEP-1999 JP 2000572239  
**PR** 29-SEP-1998 US 09/163162, 05-APR-1999 US 09/286407 PI  
**FR** FRANK C BENNETT, ELIZABETH J ACKERMANN, ERIC E SWAYZE, LEX M PI  
**CONSEPT**  
**PC** C07H21/04, A61K31/7088, A61K31/712, A61K48/00, A61P35/00 CC  
**FT** Key Location/Qualifiers  
**CD** (50)...(478).  
**FT** CDS Location/Qualifiers  
**source** 1..1619  
**FEATURES** /organism="Homo sapiens"  
**source** /mol\_type="genomic DNA"  
**ORIGIN** /db\_xref="taxon:9606"  
**Alignment Scores:**  
**Pred. No.:** 1..38e-76 Length: 1619  
**Score:** 778.00 Matches: 142  
**Percent Similarity:** 100.00% Conservative: 0  
**Best Local Similarity:** 100.00% Mismatches: 0  
**Query Match:** 100.00% Indels: 0  
**DB:** 6 Gaps: 0  
**US-09-690-825-34 (1-142) x BD273550 (1-1619)**  
**QY** 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
**Db** 50 ATGGGTGCGCGAGCTTGCCCTGCTGGCAGCCCTTCTCAAGGACCAACCGCATCTCT 109  
**QY** 21 ThrPheLysAsnTrpProPheLeuGluCysAlaCysThrProGluArgMetAlaGlu 40  
**Db** 110 ACATTCAGAACTGGCCCTTCTTGAGGGCTGCGCTGCACCCCGAGCGATGGCCGAG 169  
**QY** 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
**Db** 170 GCTGCTTCATCCACTGCCCTGACGAGCCAGCTTGCCCGAGTGTCTTCTCTGC 229  
**QY** 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
**Db** 230 TTCAAGGAGCTGGAGGCTGGAGCCAGATGACGCCATAGAGGAACATAAAAGCAT 289  
**QY** 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
**Db** 290 TCGTCCGGTTGCGCTTCTTCTGTCAGAGGAGCATTTGAAGATTAACCTTTGGTGAA 349  
**QY** 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLysLysGluThrAsnAsnLys 120  
**Db** 350 TTTTGTAACTGGACAGAGAAAGAGCCAAAGAAACAAAATTCAGAGGAACCAACAATAAG 409  
**QY** 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140  
**Db** 410 AAGAAAGATTTGAGAACTGCCAAGAACTGCCCGTCCATCGAGCAGCTGGCTGCC 469  
**QY** 141 MetAsp 142  
**Db** 470 ATGGAT 475  
**RESULT 7**  
**AX775129** 1619 bp DNA linear PAT 09-JUL-2003  
**LOCUS** Sequence 445 from Patent WO03038129.  
**DEFINITION** AX775129  
**ACCESSION** AX775129  
**VERSION** AX775129.1 GI:32486645  
**KEYWORDS** Homo sapiens (human)  
**SOURCE** Homo sapiens  
**ORGANISM** Homo sapiens  
**REFERENCE** 1  
**AUTHORS** Raponi, M.  
**TITLE** Methods for assessing and treating leukemia  
**US-09-690-825-34 (1-142) x AR181635 (1-1619)**  
**QY** 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
**Db** 50 ATGGGTGCGCGAGCTTGCCCTGCTGGCAGCCCTTCTCAAGGACCAACCGCATCTCT 109  
**QY** 21 ThrPheLysAsnTrpProPheLeuGluCysAlaCysThrProGluArgMetAlaGlu 40  
**Db** 110 ACATTCAGAACTGGCCCTTCTTGAGGGCTGCGCTGCACCCCGAGCGATGGCCGAG 169  
**QY** 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
**Db** 170 GCTGCTTCATCCACTGCCCTGACGAGCCAGCTTGCCCGAGTGTCTTCTCTGC 229  
**QY** 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
**Db** 230 TTCAAGGAGCTGGAGGCTGGAGCCAGATGACGCCATAGAGGAACATAAAAGCAT 289  
**QY** 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
**Db** 290 TCGTCCGGTTGCGCTTCTTCTGTCAGAGGAGCATTTGAAGATTAACCTTTGGTGAA 349  
**QY** 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLysLysGluThrAsnAsnLys 120  
**Db** 350 TTTTGTAACTGGACAGAGAAAGAGCCAAAGAAACAAAATTCAGAGGAACCAACAATAAG 409  
**QY** 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140  
**Db** 410 AAGAAAGATTTGAGAACTGCCAAGAACTGCCCGTCCATCGAGCAGCTGGCTGCC 469  
**QY** 141 MetAsp 142  
**Db** 470 ATGGAT 475  
**RESULT 6**  
**AR181635** 1619 bp DNA linear PAT 20-APR-2002  
**LOCUS** Sequence 445 from Patent WO03038129.  
**DEFINITION** AR181635  
**ACCESSION** AR181635  
**VERSION** AR181635.1 GI:32486645  
**KEYWORDS** Homo sapiens (human)  
**SOURCE** Homo sapiens  
**ORGANISM** Homo sapiens  
**REFERENCE** 1  
**AUTHORS** Raponi, M.  
**TITLE** Methods for assessing and treating leukemia



JOURNAL Patent: WO 03038129-A 445 08-MAY-2003;  
Ortho-Clinical Diagnostics, Inc. (US)  
FEATURES  
source  
1..1619  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Alignment Scores:  
Pred. No.: 1 38e-76 Length: 1619  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x AX775129 (1-1619)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 50 ATGGGTGCCCGGACGTTGGCCCTGCTGGCAGCCCTTTCTCAAGGACCCGCGCATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 110 ACATTCAAGAACTGGCCCTTCTTGGAGGCTGGCTGCACCCCGAGCGGATGCGCGAG 169  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
DB 170 GCTGGCTTCATCCACTGCCCGGACGAGACGAGCAGACTTGGCCCGAGTGTTCCTCTGC 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluHisLysLysHis 80  
DB 230 TTCAGGAGCTGGAGGCTGGGAGCGAGATGACGCCCATAGAGAAATATAAAGCAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 290 TCGTCCGGTTGGCTTCTCTCTGCAAGAGCAGTTTGAAGATTAAACCTTGTGTAA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
DB 350 TTTTGAAGAACTGGACAGAGAAAGAGCAAGAAACAAAATTGCAAGGAAACCAACAAATAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
DB 410 AAGAAAGAAATTGAGGAAACTGCGAAGAAAGTGGCCCGTGCATCGAGCAGCTGCTGCC 469  
QY 141 MetAsp 142  
DB 470 ATGGAT 475

## RESULT 8

AX779941 1619 bp DNA linear PAT 14-JUL-2003  
LOCUS  
DEFINITION Sequence 2098 from Patent WO03039443.  
ACCESSION AX779941  
VERSION AX779941.1 GI:32696935

KEYWORDS Homo sapiens (human)

## SOURCE

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1  
AUTHORS Haferlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,  
Dugas, M., Eils, R., Brors, B. and Mergenthaler, S.  
TITLE Novel genetic markers for leukemias  
JOURNAL Patent: WO 03039443-A 2098 15-MAY-2003;  
Deutsches Krebsforschungszentrum (DE) ;  
Ludwig-Maximilians-Universität München (DE) ; Haferlach, Torsten,  
PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)

## FEATURES

source  
1..1619  
/organism="Homo sapiens"

## ORIGIN

Alignment Scores:  
Pred. No.: 1 38e-76 Length: 1619  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x AX779941 (1-1619)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 50 ATGGGTGCCCGGACGTTGGCCCTGCTGGCAGCCCTTTCTCAAGGACCCGCGCATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 110 ACATTCAAGAACTGGCCCTTCTTGGAGGCTGGCTGCACCCCGAGCGGATGCGCGAG 169  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
DB 170 GCTGGCTTCATCCACTGCCCGGACGAGACGAGCAGACTTGGCCCGAGTGTTCCTCTGC 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluHisLysLysHis 80  
DB 230 TTCAGGAGCTGGAGGCTGGGAGCGAGATGACGCCCATAGAGAAATATAAAGCAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 290 TCGTCCGGTTGGCTTCTCTCTGCAAGAGCAGTTTGAAGATTAAACCTTGTGTAA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
DB 350 TTTTGAAGAACTGGACAGAGAAAGAGCAAGAAACAAAATTGCAAGGAAACCAACAAATAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
DB 410 AAGAAAGAAATTGAGGAAACTGCGAAGAAAGTGGCCCGTGCATCGAGCAGCTGCTGCC 469  
QY 141 MetAsp 142  
DB 470 ATGGAT 475

## RESULT 9

AF077350 1629 bp mRNA linear PRI 14-DEC-2000  
LOCUS  
DEFINITION Homo sapiens inhibitor of apoptosis homolog mRNA, complete cds.  
ACCESSION AF077350  
VERSION AF077350.1 GI:4959078

KEYWORDS Homo sapiens (human)

## SOURCE

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 1629)  
AUTHORS Uren, A.G., Wong, L., Pakusch, M., Fowler, K.J., Burrows, F.J.,  
Vaux, D.L. and Choo, K.H.  
TITLE Survivin and the inner centromere protein INCENP show similar  
cell-cycle localization and gene knockout phenotype  
JOURNAL Curr. Biol. 10 (21), 1319-1328 (2000)

## MEDLINE

PUBMED 11084331

## REFERENCE

2 (bases 1 to 1629)

Uren, A.G. and Vaux, D.L.

## AUTHORS

Direct Submission  
TITLE Submitted (02-OCT-1997) Molecular Cancer Division, The Walter and  
Eliza Hall Institute, Royal Parade, Parkville, Vic 3052, Australia

## FEATURES

Location/Qualifiers  
1..1629  
/organism="Homo sapiens"

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

```

/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE consortium clone ID 590560"
/tissue_type="pancreatic adenocarcinoma"
25. .453
/notes="IAP; contains single baculovirus IAP repeat (BIR) motif"
/codon_start=1
/product="inhibitor of apoptosis homolog"
/protein_id="AAD34226.1"
/db_xref="GI:4959079"
/translation="MGAPTLPPAWQPFKDHRISTFKNWPFLGCGACTPERMAEAGFI
HCPTEPNDIAQCFCEKLEGEPPDDPIEHHKHSGCCAFUSVKQKFEELTLGFFL
KLDREARKNKIATETNNKKKEFEETAKKVRRAIEQLAAMD"

CDS
ORIGIN
Alignment Scores:
Pred. No.: 1.39e-76 Length: 1629
Score: 778.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x AF077350 (1-1629)
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
DB 25 ATGGTGCCCGCGAGTGGCCCTGCTGGCAGCCCTTCTCAAGGACCAACCGCATCTCT 84
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 85 ACATTCAGAACTGGCCCTCTTGGAGGGCTGGCCCTGCACCCCGGAGCGATGGCGGAG 144
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60
DB 145 GCTGGCTTCATCCACTGCCCTGAGAAAGAGCCAGACTTGCCCGAGTGTTCCTCTGCG 204
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80
DB 205 TTCAAGGAGCTGGAGGCTGGAGCCAGATGACGACCCCATAGAGGAACATAAAGCAT 264
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB 265 TCGTCGGGTGCGCTTTCCTTCCTGCAAGAGCAGTTTGAAGAAATTAACCTTGTGTGA 324
QY 101 PheLeuLysLeuAspArgGluAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
DB 325 TTTTGAACCTGGACAGAAAGAGCCAGCAAAATTCGAAAGGAACCAACAATAAG 384
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaIa 140
DB 385 AAGAAAGAAATTGAGAAATGCGAAGAAATGCGCGCTGCCATCGCAGCAGCTGGCTGCC 444
QY 141 MetAsp 142
DB 445 ATGGAT 450

RESULT 10
BC034148 1643 bp mRNA linear PRI 12-NOV-2003
LOCUS Homo sapiens baculoviral IAP repeat-containing 5 (survivin), mRNA
DEFINITION (CDNA clone MGC:32768 IMAGE:4656567), complete cds.
ACCESSION BC034148
VERSION BC034148.1 GI:21707886
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1643)
AUTHORS Srausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altechul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F.,
Diatchenko,L., Marusina,K., Farmer,A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.D.,
Scheetz,T.E., Brownstein,M.J., Udwin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fanej,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.C., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Kzywinski,M.I., Skalska,U., Smalilus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1643)
Strausberg,R.
Direct Submission
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
JOURNAL
TITLE
REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IPAL Plate: 41 Row: h Column: 21.
Location/Qualifiers
1. .1643
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="MGC:32768 IMAGE:4656567"
/tissue_type="lung, mucopidermoid carcinoma"
/clone_lib="NIH_MGC_59"
/lab_host="DH10B"
/notes="vector: pDNR-LIB"
1. .1643
/gene="BIRC5"
/notes="synonyms: API4, SURVIVIN, EPR-1"
/db_xref="LocusID:332"
/db_xref="MIM:603352"
45. .473
/codon_start=1
/product="baculoviral IAP repeat-containing protein 5"
/protein_id="AAH34148.1"
/db_xref="GI:21707887"
/db_xref="LocusID:332"
/translation="MGAPTLPPAWQPFKDHRISTFKNWPFLGCGACTPERMAEAGFI
HCPTEPNDIAQCFCEKLEGEPPDDPIEHHKHSGCCAFUSVKQKFEELTLGFFL
KLDREARKNKIATETNNKKKEFEETAKKVRRAIEQLAAMD"
84. .305
/notes="BIR; Region: Baculoviral inhibition of apoptosis
protein repeat"
/db_xref="CDD:smart00238"

misc_feature
ORIGIN

```

## Alignment Scores:

Pred. No.: 14e-76 Length: 1643  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservativity: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x BC034148 (1-1643)

QY 1 MetGlyAlaProThrLeuProProlaTrpGlnProPheLeuLeuAspHisArgIleSer 20  
 Db 45 ATGGGTGCCCCGACGTTGCCCTTCTGTCGAGCCCTTCTCAAGGACCCAGCGCATCTCT 104  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 105 ACATTCAAGAACTGGCCCTTCTTGAGGGCTCGCGCTGCACCCGGGCGGATGCCCGAG 164  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
 Db 165 GGTGGCTTCATCCACTGCCCACTGAGAACGAGCGAGACTTGGCCAGTGTTCCTCTGCG 224  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
 Db 225 TTCAAGGAGCTGGAAGGCTGGGAGCGAGATGACGCCCATAGAGGAAACATAAAGCAT 284  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 285 TGTCCGCTGGCTTCTCTTCTGTCAGAGAGCGAGTTGAGAAATTAACCTTGGTGA 344  
 QY 101 PheLeuLysLeuAspArgGluAlaGlnAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 Db 345 TTTTGAACCTGGACAGAGAAAGAGCAAGAAACAAAATTCAGAGGAAACCAACAAATAG 404  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaLa 140  
 Db 405 AAGAAAGAAATTTGAGAAACTGGAGAAAGTGGCCGCGGATCGCATCGAGCGTGGCTGCC 464  
 QY 141 MetAsp 142  
 Db 465 ATGGAT 470

## RESULT 11

BC008718

LOCUS

DEFINITION Homo sapiens baculoviral IAP repeat-containing 5 (survivin), mRNA (cdna clone MGC:8592 IMAGE:2961114), complete cds.

ACCESSION BC008718

VERSION BC008718.2 GI:38196997

KEYWORDS MGC.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1653)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,

Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

Schuetz,T.E., Brownstein,M.J., Usdin,T.B., Tashiro,Y.,

Carninci,P., Frange,C., Rana,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

Fahey,J., Helton,E., Kettner,M., Madan,A., Young,A.C., Shcherchenko,Y.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shcherchenko,Y.,

Bouffard,G.G., Blakesley,R.W., Touchman,O.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E.,

## TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

Schmerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 22389257  
 12477932  
 2 (bases 1 to 1653)  
 Strausberg,R.  
 Direct Submission  
 Submitted (25-MAY-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NTH-MGC Project URL: http://mgc.nci.nih.gov  
 On Nov 6, 2003 this sequence version replaced gi:14250533.  
 Contact: MGC help desk  
 Email: gcgaps-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland  
 Web site: http://www.nisc.nih.gov/  
 Contact: nisc\_mgc@nhri.nih.gov  
 Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,  
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,  
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,  
 Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,  
 Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,  
 McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,  
 Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,  
 Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAL Plate: 1 Row: 1 Column: 7

This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 5453604.

Location/Qualifiers

1. 1653

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:8592 IMAGE:2961114"

/tissue\_type="Muscle, rhabdomyosarcoma"

/clone\_lib="NIH MGC 17"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

1. 1653

/gene="BIRC5"

/notes="synonyms: API4, SURVIVIN, EPR-1"

/db\_xref="LocusID:332"

/db\_xref="MIM:603352"

34. 462

/codon\_start=1

/product="baculoviral IAP repeat-containing protein 5"

/protein\_id="AA08718.1"

/db\_xref="GI:14250534"

/db\_xref="LocusID:332"

/translation="MGAPLPAPQPPFLKDHRISTFKNWPFLGCACTPERMAAGFI  
 HCTENPDLAQCFFCFCELEGGDDDDPIEHHKSSGCGAPLSVKKQPELTLGEPFL  
 KLDRERAKNKIAETNNKKEPETAKKVRAIEQLAAMD"

73. 294

/notes="BIR; Region: Baculoviral inhibition of apoptosis  
 protein repeat"

/db\_xref="CDD:smart00238"

misc\_feature

1. 41e-76 Length: 1653  
 778.00 Matches: 142  
 100.00% Conservativity: 0

ORIGIN

Alignment Scores:

Pred. No.: 1.41e-76 Length: 1653

Score: 778.00 Matches: 142

Percent Similarity: 100.00% Conservativity: 0

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x BC008718 (1-1653)

Qy 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 34 ATGGTGGCCCGAGCTGGCCCTGCTGGCAGCCCTTTCTCAAGGACCAACCGCATCTCT 93
Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 94 ACATTCAGAACTGGCCCTTCTTGAGGGCTGGCCCTGCACCCCGAGCGATGGCCGAG 153
Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db 154 GCTGGCTTCATCCACTGCCCTGCTGAGAACGAGCCAGACTTGGCCCGAGTGTTCCTCTGC 213
Qy 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80
Db 214 TTCAAGGAGCTGGAGGCTGGAGCCAGATGACGACCCCATAGAGGAACATAAAAAAGCAT 273
Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
Db 274 TCGTCCGGTTCGGCTTCTCTGTCAGAGAGCAGTTTGAGAAATTAACCTTGGTGA 333
Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
Db 334 TTTTGAACCTGGACAGAAAGAGCCCAAGAACAAAAATTCAGAGGAACCAACAATAAG 393
Qy 121 LysLysGluPheGluGluThrAlaLysValArgArgAlaIleGluGlnLeuAlaA 140
Db 394 AAGAAAGAAATTTGAGAAACTGCGAAGAAAGTGGCCGCTCCATCGAGCAGCTGGCTGCC 453
Qy 141 MetAsp 142
Db 454 ATGGAT 459

RESULT 12
AB095108 1630 bp mRNA linear MAM 27-NOV-2002
LOCUS Canis familiaris mRNA for survivin, complete cds.
DEFINITION AB095108
ACCESSION AB095108
VERSION AB095108.1 GI:24636590
KEYWORDS
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1
Inoue, C., Yamazaki, J., Kano, R. and Hasegawa, A.
Canis familiaris mRNA for survivin-protein, complete cds
Published Only in Database (2002)
2 (bases 1 to 1630)
Inoue, C., Yamazaki, J., Kano, R. and Hasegawa, A.
Direct Submission
Submitted (01-NOV-2002) Rui Kano, Nihon University School of
Veterinary Medicine, Department of Pathobiology, Kameino1866,
Fujisawa, Kanagawa 252-8510, Japan (E-mail:Kano@vms.nihon-u.ac.jp,
Tel:81-466-84-3649, Fax:81-466-84-3649)
Location/Qualifiers
1. 1630
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/cell_line="CL-1"
/cell_type="lymphocyte"
1. 154
55. 483
/product="survivin"
/protein_id="BAC22748.2"
/db_xref="GI:25553512"

FEATURES
source
/organism="Canis familiaris"
2 (bases 1 to 600)
Kageyama, H., Islam, A., Takayasu, H. and Nakagawara, A.
An isoform of Survivin (Survivin-beta) Which Has 23 Amino Acids
Insertion into the BIR Domain
Unpublished
JOURNAL
AUTHORS
AUTHORS
TITLE
JOURNAL
Submitted (11-JUN-1999) Akira Nakagawara, Chiba Cancer Center
Research Institute, Division of Biochemistry; 666-2 Nitona,
Chuch-ku, Chiba, Chiba 260-8717, Japan
(E-mail:akiranak@chiba-cc.pref.chiba.jp,
Tel:81-43-264-5431(ex.5201), Fax:81-43-265-4459)
Location/Qualifiers

/translation="MGAPTLPPAWQPFKXDRHSTRFKWPFLEGCACCTPDRMAEAGFI
HCPTENPDLAQCFKFELEGWEPDDPIEHHKSHSGCAFLSVKKQFELTLGEFL
XLDREKAKNKIATKNTNNKKFEFETAKKVRRAIEQLAAMD"
484. 1630

3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 2,99e-76 Length: 1630
Score: 775.00 Matches: 141
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.30% Mismatches: 0
Query Match: 99.61% Indels: 0
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x AB095108 (1-1630)
Qy 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 55 ATGGTGGCCCGAGCTGGCCCTGCTGGCAGCCCTTTCTCAAGGACCAACCGCATCTCT 114
Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 115 ACATTCAGAACTGGCCCTTCTTGAGGGCTGGCCCTGCACCCCGAGCGATGGCAGAG 174
Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db 175 GCGCGCTTCATCCACTGCCCTGCTGAGAACGAGCCAGCTTGGCCCGAGTGTTCCTCTGC 234
Qy 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80
Db 235 TTCAAGGAGCTGGAGGCTGGAGCCAGATGATCACCTATAGAGGAGCATAAAAACAT 294
Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
Db 295 TCATCTGGTGTGCTTCTCTTCTGTCAAGAGCAGTTTGAAGAAATTAACCTTGGTGA 354
Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
Db 355 TTTTGAACCTGGACAGAAAGAGCCCAAGAACAAAAATTCAGAGGAACCAACAATAAG 414
Qy 121 LysLysGluPheGluGluThrAlaLysValArgArgAlaIleGluGlnLeuAlaA 140
Db 415 AAGAAAGAAATTTGAGAAACTGCGAAGAAAGTGGCCGCTCCATCGAGCAGCTGGCTGCC 474
Qy 141 MetAsp 142
Db 475 ATGGAT 480

RESULT 13
AB028869 600 bp mRNA linear PRI 04-APR-2000
LOCUS Homo sapiens mRNA for survivin-beta, complete cds.
DEFINITION AB028869
ACCESSION AB028869
VERSION AB028869.1 GI:7416052
KEYWORDS survivin-beta.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Kageyama, H., Islam, A., Takayasu, H. and Nakagawara, A.
An isoform of Survivin (Survivin-beta) Which Has 23 Amino Acids
Insertion into the BIR Domain
Unpublished
JOURNAL
AUTHORS
AUTHORS
TITLE
JOURNAL
Submitted (11-JUN-1999) Akira Nakagawara, Chiba Cancer Center
Research Institute, Division of Biochemistry; 666-2 Nitona,
Chuch-ku, Chiba, Chiba 260-8717, Japan
(E-mail:akiranak@chiba-cc.pref.chiba.jp,
Tel:81-43-264-5431(ex.5201), Fax:81-43-265-4459)
Location/Qualifiers

FEATURES

```

```
source 1..500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="CHP134"
/cell_type="neuroblastoma"
27..524
/codon_start=1
/product="survivin-beta"
/protein_id="BAA93676.1"
/db_xref="GI:7416053"
/translation="MGAPTLPPAWQPFLLKDHRISTFKNWPLEGCACTPERMAEAGFI
HCPTEPEDLAQCFKFELEGWPDPIPGFTVAVACNTSTLGGGGRRITREHKK
HSSGCAFLSVKKQFEELTLGFELKLDREAKNKIAKETNNKKFEETAKKVRRAIQ
LAAMD"

ORIGIN
Alignment Scores:
Pred. No.: 1..08e-74 Length: 600
Score: 756.50 Matches: 142
Percent Similarity: 86.06% Conservative: 0
Best Local Similarity: 86.06% Mismatches: 0
Query Match: 97.24% Indels: 23
DB: 9 Gaps: 1

US-09-690-825-34 (1-142) x AB028869 (1-600)
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
DB 27 ATGGGTGCCCCGACGTTGCCCTCGCTGGAGCGCCCTTCTCAAGGACCCCGCATCTCT 86
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 87 ACATTCAAGAAGCTGCCCTTCTTGGAGGCTGCGCTGCACCCCGGAGCGGATGGCCGAG 146
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB 147 GCTGCTTCATCCACTGCCCTGCACCTGAGAACCGAGCAGACTTGGCCCACTGTTCTTCG 206
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIle-----GluGluHis 74
DB 207 TTCAGGAGCTGGAAGGCTGGAGCCAGATGACGACCCCATTTGGCCGGCAGCGTGCT 266
QY 75 -----GluGluHis 77
DB 267 TACGCTGTATACAGCACATTTTGGAGCGCCGAGCGCGCGATCACGAGAGAGGAACAT 326
QY 78 LysLysHisSerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThr 97
DB 327 AAAAAAGCATTCGTCGGTTGCGCTTCTCTCTCAAGAGCAGTTTGAAGAATTAAACC 386
QY 98 LeuGlyGluPheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThr 117
DB 387 CTGTGTGAATTTTGAAACTGGACAGAGAAAGACCAAGACAAATTTGCAAGGAAACC 446
QY 118 AsnAsnLysLysLysGluPheGluThrAlaLysLysValArgAlaIleGluGln 137
DB 447 AACAAATAAGAAAGAAATTTGAGGAAACTGCGAAGAAAGTGGCGCGTCCATCGAGCAG 506
QY 138 LeuAlaAlaMetAsp 142
DB 507 CTGGCTGCCATGGAT 521

RESULT 14
AR097642/c
LOCUS AR097642 1165 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6072028.
ACCESSION AR097642
VERSION AR097642.1 GI:12806372
KEYWORDS
SOURCE
ORGANISM
Unclassified.

REFERENCE
1 (bases 1 to 1165)
AUTHORS
Altieri,D.C.
TITLE
Diagnostic methods useful in the characterization of
lymphoproliferative disease characterized by increased EPR-1
Patent: US 6238875-A 1 29-MAY-2001;
JOURNAL
FEATURES
Location/Qualifiers
source 1..1165
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
```

```
REFERENCE 1 (bases 1 to 1165)
AUTHORS
Altieri,D.C.
TITLE
Diagnostic methods useful in the characterization of
lymphoproliferative disease characterized by increased EPR-1
Patent: US 6238875-A 1 06-JUN-2000;
JOURNAL
FEATURES
Location/Qualifiers
source 1..1165
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 2.03e-73 Length: 1165
Score: 748.00 Matches: 140
Percent Similarity: 99.30% Conservative: 1
Best Local Similarity: 98.59% Mismatches: 1
Query Match: 96.14% Indels: 1
DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x AR097642 (1-1165)
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
DB 1114 ATGGGTGCCCCGACGTTGCCCTCGCTGGAGCGCCCTTCTCAAGGACCCCGCATCTCT 1055
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 1054 ACATTCAAGAAGCTGCCCTTCTTGGAGGCTGCGC-TGCACCCCGGAGCGGATGGCCGAG 996
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB 995 GCTGCGTTTCACTGACCTGCCCTGAGAACGAGCCAGACTTGGCCCGAGTTTCTTCTGC 936
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80
DB 935 TTCAGGAGCTGGAAGGCTGGAGCGCCAGATGACGACCCCATAGAGGAACATAAAAGCAT 876
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB 875 TCGTCCGGTTGCGCTTCTCTTCTGTCAGAGCAGTTTGAAGATTAACCCCTTGGTGA 816
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
DB 815 TTTTGAATCTGACAGAGAGAGAGCCAGCAAAATTTGCAAGGAACCAACAATAAG 756
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140
DB 755 AAGAAAGAAATTTGAGGAAACTGCGAAGAAAGTGGCGCGTGCATCGAGCAGCTGGCCTCG 696
QY 141 MetAsp 142
DB 695 ATGGAT 690

RESULT 15
AR154245/c
LOCUS AR154245 1165 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6238875.
ACCESSION AR154245
VERSION AR154245.1 GI:15122298
KEYWORDS
SOURCE
ORGANISM
Unclassified.

REFERENCE
1 (bases 1 to 1165)
AUTHORS
Altieri,D.C.
TITLE
Diagnostic methods useful in the characterization of
lymphoproliferative disease characterized by increased EPR-1
Patent: US 6238875-A 1 29-MAY-2001;
JOURNAL
FEATURES
Location/Qualifiers
source 1..1165
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
```

Alignment Scores: Pred. No.: 2.03e-73 Length: 1165  
Score: 748.00 Matches: 140  
Percent Similarity: 99.30% Conservative: 1  
Best Local Similarity: 98.59% Mismatches: 1  
Query Match: 96.14% Indels: 1  
DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x AR154245 (1-1165)

Qy 1 MetGlyAlaProThrLeuProProAlaTTPGlnProPheLeuLysAspHisArgIleSer 20  
Db 1114 ATGGGTGCCCGACGTTGCCCTTGCCCTGGCAGCCCTTTCTCAAGGACCACCGCATCTCT 1055

Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 1054 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGCGC-TGCACCCCGGAGCGGATGGCCGAG 996

Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
Db 995 GCTGGCTTCATCCACTGCCCTCCCTGAGAACGAGCCAGACTTGGCCAGTGTTCCTCTGC 936

Qy 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluHisLysLysHis 80  
Db 935 TTCAGGAGCTGGAGGCTGGAGCCAGATGACGACCCCATAGAGGACATATAAAGCAT 876

Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
Db 875 TCGTCCGGTTGGCTTTCTCTTCTGTCAAGAGCAGTTTGAGAAATTAACCCCTTGGTGAA 816

Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLeuLys 120  
Db 815 TTTTGAACCTGGACAGAGAGAGCCCAAGACAAATTCGAAAGGAAACCAACATAAG 756

Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
Db 755 AAGAAGCAATTTGAGGAACTGCGAAGAAAGTGGCGCTGCCATCGAGCAGCTGGGCTCG 696

Qy 141 MetAsp 142  
Db 695 ATGGAT 690

Search completed: August 15, 2004, 23:26:24  
Job time : 3690 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 15, 2004, 21:39:58 ; Search time 95 Seconds  
(without alignments)  
829.506 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 778

Sequence: 1 MGAPLPPAWQPFLLKDRHS.....EFEETAKKVRRAIQLAAMD 142

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09690825/runat\_11082004\_140952\_13557/app\_query.fasta\_1.327  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFCI=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09690825 @CGN 1 1 69 @runat\_11082004\_140952\_13557 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*

2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*

3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*

4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*

5: /cgn2\_6/ptodata/2/ina/PTUS.COMB.seq:\*

6: /cgn2\_6/ptodata/2/ina/backfiles.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	100.0	426	4	US-09-283-144-2
2	778	100.0	1619	3	US-09-163-162-1
3	778	100.0	1619	3	US-09-286-407-1
4	778	100.0	1619	4	US-09-496-694B-97
c 5	748	96.1	1165	3	US-08-448-722A-1
c 6	748	96.1	1165	3	US-08-189-309B-1
7	721	92.7	740	4	US-09-016-434-918
8	660	84.8	955	4	US-09-496-694B-10
9	382	49.1	14796	3	US-08-978-080-35
10	382	49.1	14796	3	US-09-630-706-10
11	382	49.1	14796	4	US-09-496-694B-3
12	347	44.6	332	4	US-09-833-381-1331

13	246.5	31.7	624	4	US-09-738-946-3	Sequence 3, Appli
14	221	28.4	417	4	US-09-283-144-1	Sequence 1, Appli
15	155	19.9	5366	4	US-09-705-872-4	Sequence 4, Appli
16	155	19.9	5502	3	US-08-836-134-1	Sequence 1, Appli
17	155	19.9	5502	4	US-09-493-784-1	Sequence 1, Appli
18	155	19.9	5502	4	US-09-023-655-1459	Sequence 1459, Ap
19	155	19.9	5984	4	US-09-705-872-2	Sequence 2, Appli
20	136.5	17.5	1134	4	US-09-502-528-2	Sequence 2, Appli
21	136.5	17.5	1337	4	US-08-127-928-1	Sequence 1, Appli
22	136.5	17.5	1739	4	US-09-502-528-1	Sequence 1, Appli
23	131	16.8	1559	4	US-09-239-867-1	Sequence 1, Appli
24	131	16.8	2580	2	US-08-511-485-7	Sequence 7, Appli
25	131	16.8	2580	4	US-09-201-936-7	Sequence 7, Appli
26	131	16.8	2580	4	US-09-011-356-7	Sequence 7, Appli
27	131	16.8	2580	4	US-09-672-717-222	Sequence 222, App
28	131	16.8	2589	3	US-08-569-749-1	Sequence 1, Appli
29	131	16.8	2589	5	PCT-US96-12860-1	Sequence 1, Appli
30	131	16.8	3532	2	US-09-205-204-1	Sequence 1, Appli
31	131	16.8	3732	3	US-09-212-971-7	Sequence 7, Appli
32	131	16.8	3732	3	US-08-800-929A-7	Sequence 7, Appli
33	131	16.8	3732	4	US-09-617-053A-7	Sequence 3, Appli
34	130	16.7	711	3	US-09-121-979-3	Sequence 3, Appli
35	130	16.7	711	3	US-08-332-319-3	Sequence 3, Appli
36	129.5	16.6	2100	2	US-08-511-485-9	Sequence 9, Appli
37	129.5	16.6	2100	4	US-09-201-936-9	Sequence 9, Appli
38	129.5	16.6	2100	4	US-09-011-356-9	Sequence 9, Appli
39	129.5	16.6	2100	4	US-09-672-717-224	Sequence 224, App
40	129.5	16.6	2691	3	US-09-212-971-9	Sequence 9, Appli
41	129.5	16.6	2691	3	US-08-800-929A-9	Sequence 9, Appli
42	129.5	16.6	2691	4	US-08-617-053A-9	Sequence 9, Appli
43	129.5	16.6	2862	3	US-08-569-749-13	Sequence 13, Appli
44	129.5	16.6	2862	5	PCT-US96-12860-13	Sequence 13, Appli
45	129.5	16.6	3151	3	US-09-212-971-13	Sequence 13, Appli

#### ALIGNMENTS

RESULT 1  
US-09-283-144-2  
; Sequence 2, Application US/09283144  
; Patent No. 6346389  
; GENERAL INFORMATION:  
; APPLICANT: Yale University  
; TITLE OF INVENTION: Method for Selectively Modulating the Interactions  
; FILE REFERENCE: between Survivin and Tubulin  
; CURRENT APPLICATION NUMBER: US/09/283,144  
; CURRENT FILING DATE: 1999-04-01  
; EARLIER APPLICATION NUMBER: US 60/080,288  
; EARLIER FILING DATE: 1998-04-01  
; EARLIER APPLICATION NUMBER: US 08/975,080  
; EARLIER FILING DATE: 1997-11-20  
; EARLIER APPLICATION NUMBER: PCT/US97/21880  
; EARLIER FILING DATE: 1997-11-20  
; EARLIER APPLICATION NUMBER: US 60/031,435  
; EARLIER FILING DATE: 1996-11-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 426  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Nucleotide sequence of open reading frame encoding  
; OTHER INFORMATION: Survivin  
US-09-283-144-2

Alignment Scores:  
Pred. No.: 3 63e-93  
Score: 778.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
Length: 426  
Matches: 142  
Conservative: 0  
Mismatches: 0  
Indels: 0

```

DB: 4 Gaps: 0
US-09-690-825-34 (1-142) x US-09-283-144-2 (1-426)

Qy 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 1 ATGGGTGCCCGAGCTTGGCCCTTGGCCCTGCTGGCAGCCCTTCTCAAGAGACCACCGCATCTCT 60

Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 61 ACATTCAGAACTGGCCCTTCTTGGAGGCTGGCCCTGCTGGCAGCCGATGGCCGAG 120

Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db 121 GCTGGCTTCATCCACTGCCCTTCTTGGAGGCTGGCCCTGCTGGCAGCCGATGGCCGAG 180

Qy 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluHisLysLysHis 80
Db 181 TTCAAGAGCTGGAGGCTGGAGCCAGATGACGCCATAGAGGAACATAAAGCAT 240

Qy 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluLeuThrLeuGlyGlu 100
Db 241 TCGTCCGGTTGGCTTCTTCTGTCAGAGAGCAGTTTGAAGAAATTAACCTTGGTGAA 300

Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
Db 301 TTTTGAACCTGGACAGAGAGGCGCAAGACAAATTTGCAAGAGAAACCAACAATAAG 360

Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140
Db 361 AAGAAAGAAATTTGAGGAACTGCGAGAAAGTGGCCGTCGATCGAGCAGCTGGCTGCC 420

Qy 141 MetAsp 142
Db 421 ATGGAT 426

RESULT 2
US-09-163-162-1
; Sequence 1, Application US/09163162
; Patent No. 6077709
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Ackermann, Elizabeth J.
; APPLICANT: Swayze, Eric B.
; APPLICANT: Cowsett, Lex M.
; TITLE OF INVENTION: ANTISENSE MODULATION OF Survivin EXPRESSION
; FILE REFERENCE: RTS-0008
; CURRENT APPLICATION NUMBER: US/09/163,162
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)..(478)
US-09-163-162-1

Alignment Scores:
Pred. No.: 2,7e-92 Length: 1619
Score: 778.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-163-162-1 (1-1619)

Qy 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 50 ATGGGTGCCCGAGCTTGGCCCTTGGCCCTGCTGGCAGCCCTTCTCAAGAGACCACCGCATCTCT 109

Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 110 ACATTCAGAACTGGCCCTTCTTGGAGGCTGGCCCTGCTGGCAGCCGATGGCCGAG 169

Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db 170 GCTGGCTTCATCCACTGCCCTTCTTGGAGGCTGGCCCTGCTGGCAGCCGATGGCCGAG 229

```

```

Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 110 ACATTCAGAACTGGCCCTTCTTGGAGGCTGGCCCTGCTGGCAGCCGATGGCCGAG 169

Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db 170 GCTGGCTTCATCCACTGCCCTTCTTGGAGGCTGGCCCTGCTGGCAGCCGATGGCCGAG 229

Qy 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluHisLysLysHis 80
Db 230 TTCAAGAGCTGGAGGCTGGAGCCAGATGACGCCATAGAGGAACATAAAGCAT 289

Qy 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluGluThrLeuGlyGlu 100
Db 290 TCGTCCGGTTGGCTTCTTCTGTCAGAGAGCAGTTTGAAGAAATTAACCTTGGTGAA 349

Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
Db 350 TTTTGAACCTGGACAGAGAGGCGCAAGACAAATTTGCAAGAGAAACCAACAATAAG 409

Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140
Db 410 AAGAAAGAAATTTGAGGAACTGCGAGAAAGTGGCCGTCGATCGAGCAGCTGGCTGCC 469

Qy 141 MetAsp 142
Db 470 ATGGAT 475

RESULT 3
US-09-286-407-1
; Sequence 1, Application US/09286407A
; Patent No. 6165788
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Ackermann, Elizabeth J.
; APPLICANT: Swayze, Eric B.
; APPLICANT: Cowsett, Lex M.
; TITLE OF INVENTION: ANTISENSE MODULATION OF Survivin EXPRESSION
; FILE REFERENCE: ISPH-0349
; CURRENT APPLICATION NUMBER: US/09/286,407A
; CURRENT FILING DATE: 1999-04-05
; NUMBER OF SEQ ID NOS: 48
; SEQ ID NO 1
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)..(478)
US-09-286-407-1

Alignment Scores:
Pred. No.: 2,7e-92 Length: 1619
Score: 778.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-286-407-1 (1-1619)

Qy 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 50 ATGGGTGCCCGAGCTTGGCCCTTGGCCCTGCTGGCAGCCCTTCTCAAGAGACCACCGCATCTCT 109

Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 110 ACATTCAGAACTGGCCCTTCTTGGAGGCTGGCCCTGCTGGCAGCCGATGGCCGAG 169

Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db 170 GCTGGCTTCATCCACTGCCCTTCTTGGAGGCTGGCCCTGCTGGCAGCCGATGGCCGAG 229

```



QY 61 PheLysGluLeuGluGlyTrpGluProAspAspPheGluGluHisLysHis 80  
DB 230 TTCAAGGAGCTGGAAGGCTGGAGCCAGATACGACCCCATAGAGGACATAAAAGCAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 290 TCGTCGGGTGGCTTTCTCTCTGTCAGAGAGCGATTGAGAAATTAACCCCTTGGTGA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLysLysLysLysLys 120  
DB 350 TTTTGAAGAACTGGACAGAGAGAGCCCAAGAAATTTGCAAGGAAACCAACAATAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaLysLysLysLysLysLys 140  
DB 410 AGAAGAAGATTGAGAAACTGCGAAGAAAGTGGCCGCTGCATCGACGCTGGCTGCC 459  
QY 141 MetAsp 142  
DB 470 ATGGAT 475  
RESULT 4  
US-09-496-694B-97  
; Sequence 97, Application US/09496694B  
; Patent No. 6335194  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Elizabeth J. Ackermann  
; APPLICANT: Eric E. Swazy  
; APPLICANT: Lex M. Cowsert  
; FILE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
; FILE REFERENCE: ISPH-0439  
; CURRENT APPLICATION NUMBER: US/09/496,694B  
; PRIOR FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 09/286,407  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 09/163,162  
; PRIOR FILING DATE: 1998-03-29  
; NUMBER OF SEQ ID NOS: 249  
; SEQ ID NO 97  
; LENGTH: 1619  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-496-694B-97

Alignment Scores:  
Pred. No.: 2,7e-92 Length: 1619  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-496-694B-97 (1-1619)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 50 ATGGGTGCCCGACGCTGGCCCTGCCCTGGAGCCCTTTCTCAAGGACCCACCGCATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 110 ACATTCAAGAACTGGCCCTTCTTGGAGGCTCGGCTGACCCCGAGCGGATGGCCGAG 169  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 170 GCTGGCTTCATCCACTGCCCCACTGAGAACGAGCAGACTTGGCCCGAGTGTCTCTTCG 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspPheGluGluHisLysLysHis 80  
DB 230 TTCAAGGAGCTGGAAGGCTGGAGCCAGATACGACCCCATAGAGGACATAAAAGCAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 290 TCGTCGGGTGGCTTTCTCTCTGTCAGAGAGCGATTGAGAAATTAACCCCTTGGTGA 349

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLysLysLysLysLys 120  
DB 350 TTTTGAAGAACTGGACAGAGAGAGCCCAAGAAATTTGCAAGGAAACCAACAATAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaLysLysLysLysLysLys 140  
DB 410 AGAAGAAGATTGAGAAACTGCGAAGAAAGTGGCCGCTGCATCGACGCTGGCTGCC 459  
QY 141 MetAsp 142  
DB 470 ATGGAT 475

## RESULT 5

US-08-448-722A-1/c  
; Sequence 1, Application US/08448722A  
; Patent No. 6072028  
; GENERAL INFORMATION:  
; APPLICANT: Altieri, Dario C.  
; TITLE OF INVENTION: No. 6072028el Cell Surface Receptor, Antibody  
; TITLE OF INVENTION: Compositions, and Methods of Using Same  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Office of Patent Counsel, The Scripps  
; ADDRESSEE: Research Institute  
; STREET: 10550 No. 6072028th Torrey Pines Road, TPC-8  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,722A  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/189,309  
FILING DATE: 28-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: 233.1 Div1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 784-2937  
TELEFAX: (619) 784-9399  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1165 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 37..1047

US-08-448-722A-1  
Alignment Scores:  
Pred. No.: 1.44e-88 Length: 1165  
Score: 748.00 Matches: 140  
Percent Similarity: 99.30% Conservative: 1  
Best Local Similarity: 98.59% Mismatches: 1  
Query Match: 96.14% Indels: 1  
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-08-448-722A-1 (1-1165)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 1114 ATGGGTGCCCGAGTTCGCCCTGCTGCGAGCCCTTTCTCAAGGACCAACCGCATCTCT 1055  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 1054 ACATTCAAGAACTGGCCCTTTCTGGAGGCTGGCG-TGCACCCCGGAGCGGATGGCCGAG 996  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 Db 995 GCTGGCTTCATCCACTGCCCACTGAGAACGACGACTTGCCCGCAGTTCCTTCTCTCG 936  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluHisLysLysHis 80  
 Db 935 TTCAAGGAGCTGGAAGCTGGAGCCAGATGACGACCCCAATAGAGAACTATAAAAGCAT 876  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 875 TCGTCCGGTTCGCTTCCTTTCTGTCAAGAGCAGTTGAAGAAATTAACCCCTTGGTGA 816  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLysLysLysLysLys 120  
 Db 815 TTTTGAAGAACTGGAG 756  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAla 140  
 Db 755 AAGAAAGAAATTTAGGAAACTGCGAAGAAAGTGGCGCTGCATCAGAGAGCTGGCCCTCG 696  
 QY 141 MetAsp 142  
 Db 695 ATGGAT 690

## RESULT 6

US-08-189-309B-1/c  
 ; Sequence 1, Application US/08189309B  
 ; Patent No. 6238875  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Altieri, Dario C.  
 ; TITLE OF INVENTION: Diagnostic Methods Useful in the Characterization of  
 ; TITLE OF INVENTION: Lymphoproliferative Disease Characterized by Increased EPR-1  
 ; NUMBER OF SEQUENCES: 5  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Office of Patent Counsel, The Scripps  
 ; ADDRESSEE: Research Institute  
 ; STREET: 10550 No. 6238875th Torrey Pines Road, TPC-8  
 ; CITY: La Jolla  
 ; STATE: CA  
 ; COUNTRY: US  
 ; ZIP: 92037  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/189,309B  
 ; FILING DATE: 28-JAN-1994  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/988,897  
 ; FILING DATE: 10-DEC-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/667,957  
 ; FILING DATE: 12-MAR-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fitting, Thomas  
 ; REGISTRATION NUMBER: 34,163  
 ; REFERENCE/DOCKET NUMBER: 233.1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (619) 784-2937  
 ; TELEFAX: (619) 784-9399  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:

LENGTH: 1165 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 37..1047  
 US-08-189-309B-1

Alignment Scores:  
 Pred. No.: 1,44e-88 Length: 1165  
 Score: 748.00 Matches: 140  
 Percent Similarity: 99.30% Conservative: 1  
 Best Local Similarity: 98.59% Mismatches: 1  
 Query Match: 96.14% Indels: 1  
 DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-08-189-309B-1 (1-1165)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 1114 ATGGGTGCCCGAGTTCGCCCTGCTGCGAGCCCTTTCTCAAGGACCAACCGCATCTCT 1055  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 1054 ACATTCAAGAACTGGCCCTTTCTGGAGGCTGGCG-TGCACCCCGGAGCGGATGGCCGAG 996  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 Db 995 GCTGGCTTCATCCACTGCCCACTGAGAACGACGACTTGCCCGCAGTTCCTTCTCTCG 936  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluHisLysLysHis 80  
 Db 935 TTCAAGGAGCTGGAAGCTGGAGCCAGATGACGACCCCAATAGAGAACTATAAAAGCAT 876  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 875 TCGTCCGGTTCGCTTCCTTTCTGTCAAGAGCAGTTGAAGAAATTAACCCCTTGGTGA 816  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLysLysLysLysLys 120  
 Db 815 TTTTGAAGAACTGGAG 756  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAla 140  
 Db 755 AAGAAAGAAATTTAGGAAACTGCGAAGAAAGTGGCGCTGCATCAGAGAGCTGGCCCTCG 696  
 QY 141 MetAsp 142  
 Db 695 ATGGAT 690

## RESULT 7

US-09-016-434-918  
 ; Sequence 918, Application US/09016434  
 ; Patent No. 6500938  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Janice Au-Young  
 ; APPLICANT: Jeffrey J. Seilhamer  
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
 ; NUMBER OF SEQUENCES: 1490  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 ; STREET: 3174 PORTER DRIVE  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,434  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 918:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 740 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRAITUT01  
CLONE: 752848  
US-09-016-434-918

Alignment Scores:  
Pred. No.: 2,578-85 Length: 740  
Score: 721.00 Matches: 135  
Percent Similarity: 95.74% Conservative: 0  
Best Local Similarity: 95.74% Mismatches: 6  
Query Match: 92.67% Indels: 1  
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-016-434-918 (1-740)

Qy 2 GlyAlaProThriLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSerThr 21  
Db 22 GGGTGGCNCNCGTGGCACT-GCCTGGCNCCTTCTCAAGGACACCGCATCTCTACA 80  
Qy 22 PheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGluAla 41  
Db 81 TTCAGAACTGGCCCTTCTTGGAGGGCTGGCCTGCACCCCGAGCGGATGCCGAGGCT 140  
Qy 42 GlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCysPhe 61  
Db 141 GGCCTTCATCCACTGCCCCACTGANAACAGCAGCAGACTTGGCCCGAGTGTCTCTGCTTC 200  
Qy 62 LysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHisSer 81  
Db 201 AAGGAGCTGGAAGGCTGGGAGCCAGATGACGCCCCATGAGAACATTAAGAAGCATTCG 260  
Qy 82 SerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGluPhe 101  
Db 261 TCCGGTTGGCTTCTCTCTGTCAGAGAGCAGTTTGAAGATTAAACCTTGGTGAATT 320  
Qy 102 LeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLysLys 121  
Db 321 TTGAACCTGGACAGAAAGAGCCAAAGAACAAATTTGCAAGAAACCAACCAATTAAGAG 380  
Qy 122 LysGluPheGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaMet 141  
Db 381 AAGAATTTGAGAAACTGGAAGAAAGTGGCCCGTGCATCGAGCAGCTGGCTGCCATG 440  
Qy 142 Asp 142  
Db 441 GAT 443

RESULT 8

US-09-496-694B-10

; Sequence 10, Application US/09496694B  
; Patent No. 6335194  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Elizabeth J. Ackermann  
; APPLICANT: Eric B. Swayze  
; APPLICANT: Lex M. Cowsett  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
; FILE REFERENCE: ISPH-0439  
; CURRENT APPLICATION NUMBER: US/09/496,694B  
; CURRENT FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 09/286,407  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 09/163,162  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 249  
; SEQ ID NO 10  
; LENGTH: 955  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURES:  
; NAME/KEY: CDS  
; LOCATION: (109)...(531)  
US-09-496-694B-10

Alignment Scores:  
Pred. No.: 3,9e-77 Length: 955  
Score: 660.00 Matches: 118  
Percent Similarity: 92.86% Conservative: 12  
Best Local Similarity: 84.29% Mismatches: 10  
Query Match: 84.83% Indels: 0  
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-496-694B-10 (1-955)

Qy 1 MetGlyAlaProThriLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 109 ATGGAGCTCCGGCGCTGCCCCAGATCTGGCAGCTGTACCTCAAGAACTACCGCATCGCC 168  
Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 169 ACCTTCAAGAACTGGCCCTTCTTGGAGGACTGGCCTGCACCCCGAGCGAATGGCGGAG 228  
Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
Db 229 GCTGGCTTCATCCACTGCCCTACCGAGACGAGCCTGATTGGCCAGTGTCTTCTGTC 288  
Qy 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
Db 289 TTTAAGAAATTGGAGGCTGGGAACCGGATGACAAACCCGATAGAGGAGCATAGAAAGCAC 348  
Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
Db 349 TCCCTCGCTGGCTTCTCTCACGTGTCAGAGAGCAGATGGAAGAACTAACCGTCAGTGA 408  
Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
Db 409 TCTTTGAACCTGGACAGACAGAGCCAAAGAACAAATTTGCAAGGAGAGCAACAAAG 468  
Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140  
Db 469 CAAAAGAGTTTGAAGAGACTGCAAGAACTACCGCTCAGTCAATTGAGCAGCTGGCTGCC 528

RESULT 9

US-08-975-080-35  
; Sequence 35, Application US/08975080  
; Patent No. 6245523  
; GENERAL INFORMATION:  
; APPLICANT: Altieri, Dario C.  
; TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS  
; TITLE OF INVENTION: CELLULAR APOPTOSIS, AND ITS MODULATION  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:



; Sequence 3, Application US/09496694B  
; Patent No. 6335194  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Elizabeth J. Ackermann  
; APPLICANT: Eric E. Swayze  
; APPLICANT: Lex M. Cowart  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
; FILE REFERENCE: ISPH-0439  
; CURRENT APPLICATION NUMBER: US/09/456,694B  
; CURRENT FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 09/286,407  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 09/163,162  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 249  
; SEQ ID NO 3  
; LENGTH: 14796  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2811)...(2921)  
; NAME/KEY: CDS  
; LOCATION: (3174)...(3283)  
; NAME/KEY: CDS  
; LOCATION: (5158)...(5275)  
; NAME/KEY: CDS  
; LOCATION: (11955)...(12044)  
US-09-496-694B-3

Alignment Scores:  
Pred. No.: 8,1e-39 Length: 14796  
Score: 382.00 Matches: 73  
Percent Similarity: 46.84% Conservative: 1  
Best Local Similarity: 46.20% Mismatches: 0  
Query Match: 49.10% Indels: 84  
DB: 4 Gaps: 1

US-09-690-825-34 (1-142) x US-09-496-694B-3 (1-14796)

Qy 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 2811 ATGGGTGCCCCGACGTGCCCCCTGCTGGCAGCCCTTCTCAAGGACACCGCATCTCT 2870  
Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArg----- 37  
Db 2871 ACATTCAAGAACTGGCCCTTCTTGGAGGCTCGCCCTGCACCCCGAGCGGGTGAGACTG 2930  
Qy 37 ----- 37  
Db 2931 CCGGGCTCTGGGTCCCCCAGCGCCGCTTGCCCTGTCCCTAGCGAGGCCACTGTGAC 2990  
Qy 37 ----- 37  
Db 2991 TGGGCTCGGGGTACAGCGCGCTCCCTCCCGCTGCTGCCAGCGAGGCCACTGT 3050  
Qy 37 ----- 37  
Db 3051 GGCTGGCCCCCTTGGGTCCAGCGCGGCTCCCTCCCTGCTTGTCCCATCGAGGCTT 3110  
Qy 37 ----- 37  
Db 3111 TGTGGCTGGGCTCGGGGTCCGGGTGCCACGTCCACTCAGAGTGTGTCCCTTG 3170  
Qy 38 ---MetAlaGluAlaGlyPheIleHisCysProThrGluAsnGluProAspLeuIaGln 56  
Db 3171 CAGATGCGCGAGCTGGCTTCATCCACTGCCCTCCCTCAGAGCGAGCTGGCCGAG 3230  
Qy 57 CysPhePheCysPheLysGluLeuGluGlyTrpGluProAspAspProIle 74  
Db 3231 TGTTTCTTCTTCAAGAGCTGGAGGCTGGAGCGGAGCGGAGTGCACCCCATG 3284

RESULT 12  
US-09-833-381-1331  
; Sequence 1331, Application US/09833381  
; Patent No. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1331  
; LENGTH: 332  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(332)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1331

Alignment Scores:  
Pred. No.: 1.06e-36 Length: 332  
Score: 347.00 Matches: 69  
Percent Similarity: 72.45% Conservative: 2  
Best Local Similarity: 70.41% Mismatches: 25  
Query Match: 44.60% Indels: 2  
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-833-381-1331 (1-332)

Qy 2 GlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSerThr 21  
Db 41 GGTGNCNGACNTTGNCCCTTGTGAGGCTCGCTGCACCCGNGGCGATGCGCGGC 100  
Qy 22 PheLys-AsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGluAl 41  
Db 101 TTCAGGAATCGACCTTCTTGGAGGCTCGCTGCACCCGNGGCGATGCGCGGC 160  
Qy 41 aGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCysPh 61  
Db 161 TGGNTTCATCCANTGNCCTGNAACGAGCCANACTTGGTCCANTGTTNTTGTCTA 220  
Qy 61 eLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 81  
Db 221 TAAGGACTGTAAAGCTGNGAGCCAGATGACGACCCCNATAGAGGAACATATAAGCA 280  
Qy 81 r-SerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThr 97  
Db 281 GTTACGGTTGGCTNNNNNTTCTGTCAAGAAACAGITTTGAAGATTAAACC 330

RESULT 13  
US-09-738-946-3  
; Sequence 3, Application US/09738946  
; Patent No. 6579701  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: DROSOPHILA HOMOLOGUES OF GENES AND PROTEINS IMPLICATED IN CANCER  
; FILE REFERENCE: EX00-043C  
; CURRENT APPLICATION NUMBER: US/09/738,946  
; CURRENT FILING DATE: 2000-12-14  
; PRIOR APPLICATION NUMBER: 60/170,832  
; PRIOR FILING DATE: 1999-12-14  
; PRIOR APPLICATION NUMBER: 60/170,838  
; PRIOR FILING DATE: 1999-12-14  
; PRIOR APPLICATION NUMBER: 60/178,580  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/185,879  
; PRIOR FILING DATE: 2000-02-29

;; PRIOR APPLICATION NUMBER: 60/185,880  
;; PRIOR FILING DATE: 2000-02-29  
;; PRIOR APPLICATION NUMBER: 60/186,150  
;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: 60/189,701  
;; PRIOR FILING DATE: 2000-03-15  
;; NUMBER OF SEQ ID NOS: 14  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 3  
;; LENGTH: 624  
;; TYPE: DNA  
;; ORGANISM: Drosophila melanogaster  
US-09-738-946-3

Alignment Scores:  
Pred. No.: 4,39e-23 Length: 624  
Score: 246.50 Matches: 48  
Percent Similarity: 57.14% Conservative: 20  
Best Local Similarity: 40.34% Mismatches: 50  
Query Match: 31.68% Indels: 1  
DB: 4 Gaps: 1

US-09-690-825-34 (1-142) x US-09-738-946-3 (1-624)

QY 14 LeuLysAspHisArgIleSerThrPheLysAsnTrpProPheLeuGluGlyCysAlaCys 33  
DB 139 CTGAAACAGCATCGCGTGGAGAGCTCAACAGATTGGCCCTTCCGGAGACCGCATCTGC 198  
QY 34 ThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrGluAsnGluProAsp 53  
DB 199 AGCATTCGAGATGGCGGAGCGGGGATTCATTGGACGGGCACCAAGCGGGAACGAC 258  
QY 54 LeuAlaGlnCysPhePheCysPheLysGluLeuGluGlyTrpGluProAspAspPro 73  
DB 259 ACTGCCACTTGTGTGTGCGGAAGACCCCTGGATGGTGGGAGCCCGAAGATGATCG 318  
QY 74 IleGluGluHisLysLysHisSerSerGlyCysAlaPheLeuSerValLysLysGlnPhe 93  
DB 319 TGGNAGAGGACGTGAACATGACCCCAATGGAGTTCGCCAAGTATCGTGCCCGAA 378  
QY 94 GluGluLeuThrLeuGluGluPheLeuLysLeuAspArgGluArgAlaLysAsnLysIle 113  
DB 379 AGGAATTAACCGTATCACAAATTCGTGAATATCTGGAATATCTGGAACCGTCGTTAAAGCGACATA 438  
QY 114 AlaLysGluThrAsnAsnLysLysLysGluPhe---GluGluThrAlaLysLysVal 131  
DB 439 GAGAAACCTGCAAGCCCTCAATCGAGCTTCGTCGGGAGAAATGAGAGCGGTCTTA 495

RESULT 14  
US-09-283-144-1  
; Sequence 1, Application US/09283144  
; Patent No. 6346389  
; GENERAL INFORMATION:  
; APPLICANT: Yale University  
; TITLE OF INVENTION: Method for Selectively Modulating the Interactions  
; TITLE OF INVENTION: between Survivin and Tubulin  
; FILE REFERENCE: 44574-5033-US  
; CURRENT APPLICATION NUMBER: US/09/283,144  
; EARLIER FILING DATE: 1999-04-01  
; EARLIER APPLICATION NUMBER: US 60/080,288  
; EARLIER FILING DATE: 1998-04-01  
; EARLIER APPLICATION NUMBER: US 08/975,080  
; EARLIER FILING DATE: 1997-11-20  
; EARLIER APPLICATION NUMBER: PCT/US97/21880  
; EARLIER FILING DATE: 1997-11-20  
; EARLIER APPLICATION NUMBER: US 60/031,435  
; EARLIER FILING DATE: 1996-11-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 417  
; TYPE: DNA  
; ORGANISM: Homo sapiens

;; FEATURE:  
;; OTHER INFORMATION: DNA sequence of proximal 5' flanking region of  
;; OTHER INFORMATION: Survivin gene  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (291)..(401)  
US-09-283-144-1

Alignment Scores:  
Pred. No.: 5,37e-20 Length: 417  
Score: 221.00 Matches: 38  
Percent Similarity: 92.86% Conservative: 1  
Best Local Similarity: 90.48% Mismatches: 3  
Query Match: 28.41% Indels: 0  
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-283-144-1 (1-417)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 231 ATGGTGGCCCGACGTTGCCCTTGGCCCTTCTCAGGACCCACCGCATCTCT 350  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 351 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGGCCCTGACCCCGAGCGGGTGAGACTG 410  
QY 41 AlaGly 42  
DB 411 CCCGGC 416

RESULT 15

US-09-705-872-4

; Sequence 4, Application US/09705872  
; Patent No. 6617429  
; GENERAL INFORMATION:  
; APPLICANT: Joh-E IKEDA  
; APPLICANT: Kenji YAMAMOTO  
; TITLE OF INVENTION: APOPTOSIS INHIBITORY PROTEIN, GENE ENCODING THE PROTEIN  
; TITLE OF INVENTION: AND CDNA THEREOF  
; FILE REFERENCE: 2000-1110/LC/00653  
; CURRENT APPLICATION NUMBER: US/09/705,872  
; CURRENT FILING DATE: 2000-11-06  
; PRIOR APPLICATION NUMBER: 09/239,797  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 5366  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (292)..(4176)  
US-09-705-872-4

Alignment Scores:

Pred. No.: 1,18e-09 Length: 5366  
Score: 155.00 Matches: 34  
Percent Similarity: 58.33% Conservative: 15  
Best Local Similarity: 40.48% Mismatches: 29  
Query Match: 19.92% Indels: 6  
DB: 4 Gaps: 3

US-09-690-825-34 (1-142) x US-09-705-872-4 (1-5366)

QY 15 LysAspHisArgIleSerThrPheLysAsnTrpProPhe---LeuGluGlyCysAlaCys 33  
DB 766 GAGGAGGCTGACTTGATCTTCAGAACTGCCATTTATGTCCAAGG-----ATA 819  
QY 34 ThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrGluAsnGluProAsp 53  
DB 820 TCCCTTGTGTGCTCTCAGAGGCTGGCTTGTCTTT-----ACAGGTAACAGGAC 870

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 15, 2004, 23:26:29 ; Search time 509 Seconds  
(without alignments)  
1368.838 Million cell updates/sec

Title: US-09-690-825-34  
Perfect score: 778  
Sequence: 1 MGATLTPAWQFLKDHRS.....EFETAKKVRRAQLAAMD 142

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -MODE=frames\_p2n.model -DEV=xlh

-Q=/cgn2\_1/USPTO\_spool/US09690845/runat\_11082004\_140953\_13643/app\_query.fasta\_1.327  
-DB=Published Applications NA -QMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=ptc -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USFR=US09690845 @CNG 1 1 520 @runat\_11082004\_140953\_13643  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGJOG -TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/1/pubpna/PTC\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/1/pubpna/PTCUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description

1	778	100.0	429	17	US-10-665-975-1	Sequence 1, Appli
2	778	100.0	1619	9	US-09-918-186A-97	Sequence 97, Appl
3	778	100.0	1619	13	US-10-342-867-566	Sequence 566, App
4	778	100.0	1619	13	US-10-172-118-566	Sequence 566, App
5	778	100.0	1619	13	US-10-181-316-97	Sequence 97, Appl
6	778	100.0	1619	13	US-10-388-360-324	Sequence 324, App
7	778	100.0	1619	17	US-10-283-975A-445	Sequence 445, App
8	778	100.0	2404	14	US-10-071-766-109	Sequence 109, App
9	778	100.0	2404	15	US-10-064-817-385	Sequence 285, App
10	721	92.7	740	16	US-10-305-720-918	Sequence 918, App
11	660	84.8	955	9	US-09-918-186A-10	Sequence 10, Appl
12	660	84.8	955	13	US-10-181-316-10	Sequence 3, Appli
13	382	49.1	14796	13	US-10-181-316-3	Sequence 973, App
14	382	49.1	14796	9	US-09-954-456-973	Sequence 1636, Ap
15	382	49.1	14796	9	US-09-954-456-1636	Sequence 3, Appli
16	382	49.1	14796	9	US-09-918-186A-3	Sequence 3421, Ap
17	382	49.1	14796	9	US-09-880-107-3421	Sequence 522, App
18	382	49.1	14796	12	US-09-968-007A-522	Sequence 35, Appl
19	382	49.1	14796	15	US-10-138-619-35	Sequence 1331, Ap
20	347	44.6	332	9	US-09-833-381-1331	Sequence 9972, Ap
21	298	38.3	555	15	US-10-029-386-9972	Sequence 33, Appl
22	233	29.9	121	15	US-10-179-730-33	Sequence 23672, A
23	216	27.8	120	15	US-10-029-386-23672	Sequence 2, Appli
24	214	27.5	1100	15	US-10-108-877-2	Sequence 1459, Ap
25	155	19.9	5502	8	US-08-913-322-1	Sequence 2, Appli
26	155	19.9	5504	8	US-10-641-643-1459	Sequence 1, Appli
27	155	19.9	5984	15	US-10-285-408-2	Sequence 2, Appli
28	155	19.9	6124	8	US-08-913-322-21	Sequence 16, Appl
29	155	19.9	6124	9	US-09-867-768A-184	Sequence 16, Appl
30	155	19.9	6133	8	US-08-913-322-2	Sequence 2, Appli
31	155	19.9	6228	8	US-08-913-322-23	Sequence 23, Appl
32	143	18.4	3773	15	US-10-041-859-1	Sequence 1, Appli
33	141	18.1	739	15	US-09-764-861-16	Sequence 16, Appl
34	141	18.1	739	13	US-09-764-861-16	Sequence 16, Appl
35	141	18.1	739	13	US-10-103-313-71	Sequence 71, Appl
36	141	18.1	739	15	US-10-115-928-16	Sequence 16, Appl
37	139.5	17.9	547	13	US-10-424-598-103666	Sequence 103666
38	138	17.7	1068	15	US-10-203-708-21	Sequence 21, Appl
39	138	17.7	5314	15	US-10-311-455-134	Sequence 134, App
40	138	17.7	5314	15	US-10-240-452-10	Sequence 10, Appl
41	137	17.6	1758	12	US-10-343-115-1	Sequence 1, Appli
42	136.5	17.5	843	15	US-10-244-586-2	Sequence 2, Appli
43	136.5	17.5	1168	16	US-10-188-646-12	Sequence 12, Appl
44	136.5	17.5	1260	16	US-10-188-646-4	Sequence 4, Appli
45	136.5	17.5	1337	15	US-10-235-026-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-10-665-975-1  
; Sequence 1, Application US/10665975  
; Publication No. US20040138119A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; APPLICANT: Tamm, Ingo  
; TITLE OF INVENTION: USE OF HEPATITIS VIRUS B X-INTERACTING  
; FILE REFERENCE: BURNHAM.005A  
; CURRENT APPLICATION NUMBER: US/10/665,975  
; CURRENT FILING DATE: 2003-09-18  
; PRIOR APPLICATION NUMBER: 60/412,109  
; PRIOR FILING DATE: 2002-09-18  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 429  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-665-975-1

Alignment Scores:

Pred. No.: 3.05e-95 Length: 429  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 17 Gaps: 0

US-09-690-825-34 (1-142) x US-10-665-975-1 (1-429)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 DB 1 ATGGGTGCCCCGAGGTGGCCCCCTGCTGGAGGCCCTTCTCAAGGACCAACCCGATCTCT 60  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 61 ACATTCAGAACTGGCCCTTCTGGAGGGCTGCGCTGCACCCCGAGCGATGGCGGAG 120  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
 DB 121 GCTGCTTCATCCACTGCCCCCACTGAGAACGAGCCAGCTGGCCCCAGTGTCTTCTGTC 180  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
 DB 181 TTCAAGGAGCTGGAGGCTGGAGCCAGATGACGACCCCACTAGAGGAACATAAAAGCAT 240  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 DB 241 TCGTCCGGTGGCGTTCCTTCTTCTGTCAGAGGAGCTTGAAGAAATTAACCTTGGTGAA 300  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 DB 301 TTTTGGAACTGGACAGAGGAGCCAGACAAATTCAGAGGAACCAACATTAAG 360  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
 DB 361 AAGAAAGAAATTTGAGAAATTCGAGAAAGTGGCGGTCATCGAGCAGCTGGCTGCC 420  
 QY 141 MetAsp 142  
 DB 421 ATGGAT 426

## RESULT 2

US-09-918-186A-97  
 ; Sequence 97, Application US/09918186A  
 ; Patent No. US20020137708A1

GENERAL INFORMATION:

APPLICANT: C. Frank Bennett  
 APPLICANT: Elizabeth J. Ackermann  
 APPLICANT: Eric B. Swayze  
 APPLICANT: Lex M. Cowsett

TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION

FILE REFERENCE: ISPH-0585

CURRENT APPLICATION NUMBER: US/09/918,186A

PRIOR FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: 09/496,694

PRIOR FILING DATE: 2000-02-02

PRIOR APPLICATION NUMBER: 09/286,407

PRIOR FILING DATE: 1999-04-05

PRIOR APPLICATION NUMBER: 09/163,162

PRIOR FILING DATE: 1998-09-29

NUMBER OF SEQ ID NOS: 250

SEQ ID NO 97

LENGTH: 1619

TYPE: DNA

ORGANISM: Homo sapiens

US-09-918-186A-97

Alignment Scores:

Pred. No.: 1.93e-94 Length: 1619  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x US-09-918-186A-97 (1-1619)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 DB 50 ATGGGTGCCCCGAGGTGGCCCCCTGCTGGAGGCCCTTCTCAAGGACCAACCCGATCTCT 109  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 110 ACATTCAGAACTGGCCCTTCTGGAGGGCTGCGCTGCACCCCGAGCGATGGCGGAG 169  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
 DB 170 GCTGCTTCATCCACTGCCCCCACTGAGAACGAGCCAGCTGGCCCCAGTGTCTTCTGTC 229  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
 DB 230 TTCAAGGAGCTGGAGGCTGGAGCCAGATGACGACCCCACTAGAGGAACATAAAAGCAT 289  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 DB 290 TCGTCCGGTGGCGTTCCTTCTGTCAGAGGAGCTTGAAGAAATTAACCTTGGTGAA 349  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 DB 350 TTTTGGAACTGGACAGAGGAGCCAGAAATTCAGAGGAACCAACATTAAG 409  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
 DB 410 AAGAAAGAAATTTGAGAAATTCGAGAAAGTGGCGGTCATCGAGCAGCTGGCTGCC 469  
 QY 141 MetAsp 142  
 DB 470 ATGGAT 475

## RESULT 3

US-10-342-887-566

; Sequence 566, Application US/10342887

; Publication No. US20040058340A1

GENERAL INFORMATION:

APPLICANT: Dai, Hongyue

APPLICANT: He, Yudong

APPLICANT: Linsley, Peter S.

APPLICANT: Mao, Mao

APPLICANT: Roberts, Christopher J.

APPLICANT: Van 't Veer, Laura Johanna

APPLICANT: Van de Vijver, Marc J.

APPLICANT: Bernards, Rene

TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

FILE REFERENCE: 9301-188-999

CURRENT APPLICATION NUMBER: US/10/342,887

CURRENT FILING DATE: 2003-01-15

PRIOR APPLICATION NUMBER: 60/298,918

PRIOR FILING DATE: 2001-06-18

PRIOR APPLICATION NUMBER: 60/380,710

PRIOR FILING DATE: 2002-05-14

PRIOR APPLICATION NUMBER: 10/172,118

PRIOR FILING DATE: 2002-06-14

NUMBER OF SEQ ID NOS: 2699

SEQ ID NO 566

LENGTH: 1619

TYPE: DNA

ORGANISM: Homo sapiens

US-10-342-887-566

Alignment Scores:

Pred. No.: 1.93e-94 Length: 1619  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0



US-09-690-825-34 (1-142) x US-10-342-887-566 (1-1619)

```
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 50 ATGGGTGCCCGACGTTGGCCCTGCTGGCAGCCCTTTCTCAAGGACCACCGCATCTCT 109
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 110 ACATTCAAGAACTGGCCCTTTCTGGAGGCTGCGCTGCACCCCGGAGCGATGGCCGAG 169
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db 170 GTGTGCTTCATCCACTGCCCACTGAGAACGACGACACTTGGCCCGATGTTCTTCTGCG 229
QY 61 PheLysGluLeuGluGlyTrpGluProAspLeuAlaGlnCysPhePheCys 80
Db 230 TTCAAGGAGCTGGAAGGCTGGGAGCGCAGATGACGACGACGACGACGACGACGACGAC 289
QY 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluGluLeuThrLeuGlyGlu 100
Db 290 TGTCCGGTTGCGCTTTCTTCTCAAGAGCGAGTTTGAAGAAATTAACCCCTTGGTGAA 349
QY 101 PheLeuLysLeuAspArgLysAlaGlnCysAsnLysIleAlaLysGluThrAsnLys 120
Db 350 TTTTGAAGCTGCAGAGAAAGACCAAGAACAAATTGCAAGGAAACCAACAATAAG 409
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAla 140
Db 410 AAGAAAGAAATTGAGGAAATTCGGAAGAAAGTGCGCGCTGCCATCGACGCTGGCTGCC 469
QY 141 MetAsp 142
Db 470 ATGGAT 475
```

## RESULT 4

US-10-172-118-566  
; Sequence 566, Application US/10172118  
; Publication No. US20030224374A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Chris  
; APPLICANT: Van 't Veer, Laura  
; APPLICANT: Van de Vijver, Marc  
; APPLICANT: Bernards, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-175-999  
; CURRENT APPLICATION NUMBER: US/10/172,118  
; CURRENT FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 60/380,770  
; PRIOR FILING DATE: 2002-05-14  
; NUMBER OF SEQ ID NOS: 2699  
; SEQ ID NO 566  
; LENGTH: 1619  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: NM\_001168  
; DATABASE ENTRY DATE: 2001-06-18  
US-10-172-118-566

Alignment Scores:  
Pred. No.: 1,93e-94 Length: 1619  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-172-118-566 (1-1619)

```
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 50 ATGGGTGCCCGACGTTGGCCCTGCTGGCAGCCCTTTCTCAAGGACCACCGCATCTCT 109
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 110 ACATTCAAGAACTGGCCCTTTCTGGAGGCTGCGCTGCACCCCGGAGCGATGGCCGAG 169
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db 170 GTGTGCTTCATCCACTGCCCACTGAGAACGACGACACTTGGCCCGATGTTCTTCTGCG 229
QY 61 PheLysGluLeuGluGlyTrpGluProAspLeuAlaGlnCysPhePheCys 80
Db 230 TTCAAGGAGCTGGAAGGCTGGGAGCGCAGATGACGACGACGACGACGACGACGACGAC 289
QY 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluGluLeuThrLeuGlyGlu 100
Db 290 TGTCCGGTTGCGCTTTCTTCTCAAGAGCAGTTTGAAGAAATTAACCCCTTGGTGAA 349
QY 101 PheLeuLysLeuAspArgLysAlaGlnCysAsnLysIleAlaLysGluThrAsnLys 120
Db 350 TTTTGAAGCTGCAGAGAAAGACCAAGAACAAATTGCAAGGAAACCAACAATAAG 409
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAla 140
Db 410 AAGAAAGAAATTGAGGAAATTCGGAAGAAAGTGCGCGCTGCCATCGACGCTGGCTGCC 469
QY 141 MetAsp 142
Db 470 ATGGAT 475
```

## RESULT 5

US-10-181-316-97  
; Sequence 97, Application US/10181316  
; Publication No. US20030211607A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Elizabeth J. Ackermann  
; APPLICANT: Eric B. Swayze  
; APPLICANT: Lex M. Cowart  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
; FILE REFERENCE: ISPH-0650  
; CURRENT APPLICATION NUMBER: US/10/181,316  
; CURRENT FILING DATE: 2002-07-16  
; PRIOR APPLICATION NUMBER: PCT/US01/02939  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 09/496,694  
; PRIOR FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 09/286,407  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 09/163,162  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 249  
; SEQ ID NO 97  
; LENGTH: 1619  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-181-316-97

Alignment Scores:  
Pred. No.: 1,93e-94 Length: 1619  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-181-316-97 (1-1619)

```
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
|||||
```

Db 50 ATGGTGCCCGGACGTTGGCCCTGCTGGCAGCCCTTTCTCAAGACCAACCGCATCTCT 109  
 Qy 21 ThrPheLysAsnTTPProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 110 ACATTCAGAACTGGCCCTCTTGGAGGCTGGCCCTGCACCCCGGAGCGGATGGCCGAG 169  
 Qy 41 AlaGlyPheLeuHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 Db 170 GCTGGCTTCATCCACTGCCCTCTTGGAGGCTGGCCCTGCACCCCGGAGCGGATGGCCGAG 229  
 Qy 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluGluHisLysLysHis 80  
 Db 230 TTCAGAGAGCTGGAGGCTGGAGCCAGATGACGCCCTATAGAGGAACATAAAAGCAT 289  
 Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 290 TCGTCCGGTTGGCTTTCTTCTGTCAAGAACGAGTTTCAAGAAATTAACCTTGGTGAA 349  
 Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 Db 350 TTTTGTGAACCTGGACAGAGAAAGCCCAAGAAATTCAGAGGAACCAACAATAAG 409  
 Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
 Db 410 AAGAAAGAAATTTGAGGAACTGGAGAAAGTGGCCGCTGCATCGAGCAGCTGGCTGCC 469  
 Qy 141 MetAsp 142  
 Db 470 ATGGAT 475

## RESULT 6

US-10-388-360-324  
 ; Sequence 324, Application US/10388360  
 ; Publication No. US200302528A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GENOMIC HEALTH  
 ; APPLICANT: Baker, Joffre B.  
 ; APPLICANT: Cronin, Maureen T.  
 ; APPLICANT: Kiefer, Michael C.  
 ; APPLICANT: Shak, Steve  
 ; APPLICANT: Walker, Michael Graham  
 ; TITLE OF INVENTION: GENE EXPRESSION PROFILING IN BIOPSED TUMOR TISSUES  
 ; FILE REFERENCE: 39740-0001US  
 ; CURRENT APPLICATION NUMBER: US/10/388,360  
 ; PRIOR FILING DATE: 2003-03-12  
 ; PRIOR APPLICATION NUMBER: US 60/412,049  
 ; PRIOR FILING DATE: 2002-09-18  
 ; PRIOR APPLICATION NUMBER: US 60/364,890  
 ; PRIOR FILING DATE: 2002-03-13  
 ; NUMBER OF SEQ ID NOS: 384  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 324  
 ; LENGTH: 1619  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-388-360-324

## Alignment Scores:

Pred. No.: 1.93e-94 Length: 1619  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-388-360-324 (1-1619)

Qy 1 MetGlyAlaProThrLeuProProAlaTTPGlnProPheLeuLysAspHisArgIleSer 20  
 Db 50 ATGGGTGCCCGGACGTTGGCCCTGCTGGCAGCCCTTTCTCAAGGACCAACCGCATCTCT 109  
 Qy 21 ThrPheLysAsnTTPProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40

Db 110 ACATTCAGAACTGGCCCTCTTCTGGAGGCTGGCCCTGCACCCCGGAGCGGATGGCCGAG 169  
 Qy 41 AlaGlyPheLeuHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 Db 170 GCTGGCTTCATCCACTGCCCTCTTGGAGGCTGGCCCTGCACCCCGGAGCGGATGGCCGAG 229  
 Qy 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluGluHisLysLysHis 80  
 Db 230 TTCAGAGAGCTGGAGGCTGGAGCCAGATGACGCCCTATAGAGGAACATAAAAGCAT 289  
 Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 290 TCGTCCGGTTGGCTTTCTTCTGTCAAGAACGAGTTTCAAGAAATTAACCTTGGTGAA 349  
 Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 Db 350 TTTTGTGAACCTGGACAGAGAAAGCCCAAGAAATTCAGAGGAACCAACAATAAG 409  
 Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
 Db 410 AAGAAAGAAATTTGAGGAACTGGAGAAAGTGGCCGCTGCATCGAGCAGCTGGCTGCC 469  
 Qy 141 MetAsp 142  
 Db 470 ATGGAT 475

## RESULT 7

US-10-283-975A-445  
 ; Sequence 445, Application US/10283975A  
 ; Publication No. US20040110792A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ortho-Clinical Diagnostics, Inc.  
 ; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia  
 ; FILE REFERENCE: CDS 233 PCT  
 ; CURRENT APPLICATION NUMBER: US/10/283,975A  
 ; PRIOR FILING DATE: 2002-10-30  
 ; PRIOR APPLICATION NUMBER: 60/340,938  
 ; PRIOR FILING DATE: 2001-10-30  
 ; PRIOR APPLICATION NUMBER: 60/338,997  
 ; PRIOR FILING DATE: 2001-10-30  
 ; PRIOR APPLICATION NUMBER: 60/340,081  
 ; PRIOR FILING DATE: 2001-10-30  
 ; PRIOR APPLICATION NUMBER: 60/341,012  
 ; PRIOR FILING DATE: 2001-10-30  
 ; NUMBER OF SEQ ID NOS: 900  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 445  
 ; LENGTH: 1619  
 ; TYPE: DNA  
 ; ORGANISM: HUMAN  
 US-10-283-975A-445

## Alignment Scores:

Pred. No.: 1.93e-94 Length: 1619  
 Score: 778.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 17 Gaps: 0

US-09-690-825-34 (1-142) x US-10-283-975A-445 (1-1619)

Qy 1 MetGlyAlaProThrLeuProProAlaTTPGlnProPheLeuLysAspHisArgIleSer 20  
 Db 50 ATGGGTGCCCGGACGTTGGCCCTGCTGGCAGCCCTTTCTCAAGGACCAACCGCATCTCT 109  
 Qy 21 ThrPheLysAsnTTPProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 110 ACATTCAGAACTGGCCCTCTTCTGGAGGCTGGCCCTGCACCCCGGAGCGGATGGCCGAG 169  
 Qy 41 AlaGlyPheLeuHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 Db 170 GCTGGCTTCATCCACTGCCCTCTTCTGGAGGCTGGCCCTGCACCCCGGAGCGGATGGCCGAG 229

QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProileGluGluHisLysLysHis 80  
DB 230 TTCAGGAGCTGGAAGCTGGAGCCAGATGACGACCCCATAGAGGAACATAAAAGCAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 290 TCGTCGGGTGGCTTTCCCTTTCTCTCAAGAGCAGCTTTGAAGAATTAAACCTTTGGTGAA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysLysLysLysLysLysLysLysLysLys 120  
DB 350 TTTTGGAACTGGACAGAGAAAGACCCAGACAAAATTGCAAGGAAACCAACATTAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaLysLysLysLysLys 140  
DB 410 AAGAAAGAAATTGAGGAAACTGCGAAGAAAGTGCGCGCTGCCATCGAGCAGCTGGCTGCC 469  
QY 141 MetAsp 142  
DB 470 ATGGAT 475  
RESULT 8  
US-10-071-766-109  
; Sequence 109, Application US/10071766  
; Publication No. US20020192678A1  
; GENERAL INFORMATION:  
; APPLICANT: Hwei-Mei Chen  
; TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE  
; FILE REFERENCE: PA-0043 US  
; CURRENT APPLICATION NUMBER: US/10/071,766  
; CURRENT FILING DATE: 2002-02-07  
; NUMBER OF SEQ ID NOS: 144  
; SOFTWARE: PERL Program  
; SEQ ID NO 109  
; LENGTH: 2404  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. US20020192678A1 251651.4  
; LOCATION: 710, 717, 724, 982-1007  
; OTHER INFORMATION: a, t, c, g, or other  
US-10-071-766-109  
Alignment Scores:  
Pred. No.: 3,35e-94 Length: 2404  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-690-825-34 (1-142) x US-10-071-766-109 (1-2404)  
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 56 ATGGGTGCCCCGACGCTGGCCCTGCTGCGAGCCCTTTCTCAAGGACCCCGCATCTCT 115  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 116 ACATTCAAGAAGCTGGCCCTTTCTTGGAGGCTGCGCTGCACCCCGAGCGGATGCCCGAG 175  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
DB 176 GCTGGCTTCATCCACTGCCCTTCTTGGAGGCTGCGCTGCACCCCGAGCGGATGCCCGAG 175  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProileGluGluHisLysLysHis 80  
DB 236 TTTCAAGGAGCTGGAAGCTGGAGCTGGAGCCAGATGACACCCCATAGAGGAACATAAAAGCAT 295  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100

DB 296 TCGTCGGGTGGCTTTCCCTTTCTCTGTCAAGAACAGCTTTGAAGAATTAAACCTTTGGTGAA 355  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysLysLysLysLysLysLysLys 120  
DB 356 TTTTGGAACTGGACAGAGAAAGACCAAAATTGCAAGGAAACCAACATAAAG 415  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaLysLysLysLysLys 140  
DB 416 AAGAAAGAAATTGAGGAAACTGCGAAGAAAGTGCGCGCTGCCATCGAGCAGCTGGCTGCC 475  
QY 141 MetAsp 142  
DB 476 ATGGAT 481  
RESULT 9  
US-10-084-817-285  
; Sequence 285, Application US/10084817  
; Publication No. US20030119009A1  
; GENERAL INFORMATION:  
; APPLICANT: Susan Stuart  
; APPLICANT: Jed G. Nuchtern  
; APPLICANT: Sharon E. Pion  
; APPLICANT: Jason M. Shohet  
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION  
; FILE REFERENCE: PA-0046 US  
; CURRENT APPLICATION NUMBER: US/10/084,817  
; CURRENT FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 60/270,784  
; PRIOR FILING DATE: 2001-02-23  
; NUMBER OF SEQ ID NOS: 365  
; SOFTWARE: PERL Program  
; SEQ ID NO 285  
; LENGTH: 2404  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. US20030119009A1 251651.4  
; NAME/KEY: unsure  
; LOCATION: 710, 717, 724  
; OTHER INFORMATION: a, t, c, g, or other  
US-10-084-817-285  
Alignment Scores:  
Pred. No.: 3,35e-94 Length: 2404  
Score: 778.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 15 Gaps: 0  
US-09-690-825-34 (1-142) x US-10-084-817-285 (1-2404)  
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 56 ATGGGTGCCCCGACGCTGGCCCTGCTGCGAGCCCTTTCTCAAGGACCCCGCATCTCT 115  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 116 ACATTCAAGAAGCTGGCCCTTTCTTGGAGGCTGCGCTGCACCCCGAGCGGATGCCCGAG 175  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
DB 176 GCTGGCTTCATCCACTGCCCTTCTTGGAGGCTGCGCTGCACCCCGAGCGGATGCCCGAG 235  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProileGluGluHisLysLysHis 80  
DB 236 TTCAGGAGCTGGAAGCTGGAGCTGGAGCCAGATGACACCCCATAGAGGAACATAAAAGCAT 295  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 296 TCGTCGGGTGGCTTTCCCTTTCTCTGTCAAGAACAGCTTTGAAGAATTAAACCTTTGGTGAA 355

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
DB 356 TTTTGAACCTGGACAGAGAAAGCCCAAGAAACAAATTCGAAGGAAACCAACAATAAG 415  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140  
DB 416 AAGAAGAAATTTGAGAAACTGCGAAGAAAGTGGCCGCTGCCATCGAGCAGCTGGCTGCC 475  
QY 141 MetAsp 142  
DB 476 ATGGAT 481  
RESULT 10  
US-10-305-720-918  
; Sequence 918, Application US/10305720  
; Publication No. US20040010136A1  
; GENERAL INFORMATION:  
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.  
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression  
; FILE REFERENCE: PA-0002-1 CON  
; CURRENT APPLICATION NUMBER: US/10/305,720  
; PRIOR FILING DATE: 2002-11-26  
; PRIOR APPLICATION NUMBER: 09/016,434  
; PRIOR FILING DATE: 1998-01-30  
; NUMBER OF SEQ ID NOS: 1490  
; SOFTWARE: PERL Program  
; SEQ ID NO 918  
; LENGTH: 740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040010136A1 752848  
US-10-305-720-918  
; OTHER INFORMATION: a, t, c, g, or other  
Alignment Scores:  
Pred. No.: 3,15e-87 Length: 740  
Score: 721.00 Matches: 135  
Percent Similarity: 95.74% Conservative: 0  
Best Local Similarity: 95.74% Mismatches: 6  
Query Match: 92.67% Indels: 1  
DB: 16 Gaps: 0  
US-09-690-825-34 (1-142) x US-10-305-720-918 (1-740)  
QY 2 GlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSerThr 21  
DB 22 GGGTGCNCGNCGTGGCCACT-GCGTGCNCTCCCTTTCTCAAGGACCACCGCATCTCTACA 80  
QY 22 PheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGluAla 41  
DB 81 TTCAAGAACTGGCCCTCTCTGGAGGCTGGCTGCTGCACCCCGAGGGATGGCCGAGGCT 140  
QY 42 GlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCysPhe 61  
DB 141 GGCTTCATCCACTGCCCTCCCTGANAAGCAGCAGACTTGGCCCGAGTGTCTCTGCTTC 200  
QY 62 LysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHisSer 81  
DB 201 AAGAGCTGAAGCTGGAGCCGACATGACACCCCATAGAGAAACATATAAGCATTCG 260  
QY 82 SerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGluPhe 101  
DB 261 TCCGGTTGGCTTCTCTCTCTGTCAGAACAGCTTTGAAGAAATTAACCTTGGTGAATTT 320  
QY 102 LeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 121  
DB 321 TTGAACTGGACAGAGAAAGCCCAAGAAATTTGCAAGGAAACCAACAATAAGAAG 380

QY 122 LysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaMet 141  
DB 381 AAGAATTTGAGAAACTGCGAAGAAAGTGGCCGCTGCCATCGAGCAGCTGGCTGCCATG 440  
QY 142 Asp 142  
DB 441 GAT 443  
RESULT 11  
US-09-918-186A-10  
; Sequence 10, Application US/09918186A  
; Patent No. US20020137708A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Elizabeth J. Ackermann  
; APPLICANT: Eric E. Swayze  
; APPLICANT: Lex M. Cowsett  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
; FILE REFERENCE: ISPH-0585  
; CURRENT APPLICATION NUMBER: US/09/918,186A  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 09/496,694  
; PRIOR FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 09/286,407  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 09/163,162  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 250  
; SEQ ID NO 10  
; LENGTH: 955  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (109)...(531)  
US-09-918-186A-10  
Alignment Scores:  
Pred. No.: 7,51e-79 Length: 955  
Score: 660.00 Matches: 118  
Percent Similarity: 92.86% Conservative: 12  
Best Local Similarity: 84.29% Mismatches: 10  
Query Match: 84.83% Indels: 0  
DB: 9 Gaps: 0  
US-09-690-825-34 (1-142) x US-09-918-186A-10 (1-955)  
QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 109 ATGGAGCTCCGGCGCTGCCAGATCTGGCAGCTGTACCTCAAGAACTACCGCATGCC 168  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 169 ACCCTCAAGAACTGGCCCTTCTCTGGAGGACTGCGCTGCACCCCGAGCGAATGGCGGAG 228  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
DB 229 GCTGCTTCATCCACTGCCCTTACCAGAACGAGCTGATTTGGCCCGAGTGTCTCTG 288  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
DB 289 TTTAAGAAATTTGAAGGCTGGGAACCCCGATGACAAACCCGATAGAGGAGCATAGAAGCAC 348  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 349 TCCCGCTGCTGCGCTTCTCTCTGCTCAAGAGCAGATGGAAGAACTAACCCGTCAGTGAA 408  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
DB 409 TTCCTGAACCTGGACACAGAGAGCCCAAGAAATTTGCAAGGAGACCAACCAACAG 468  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140  
DB 141 TTTTGAACCTGGACAGAGAAAGCCCAAGAAATTTGCAAGGAAACCAACAATAAGAAG 380

```
Db 469 CAAAAAGAGTTTGAGAGAGACTGCAAGAGACTACCCGTCAGTCAATTGAGCAGCTGGCTGCC 528
RESULT 12
US-10-181-316-10
; Sequence 10, Application US/10181316
; Publication No. US20030211607A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Elizabeth J. Ackermann
; APPLICANT: Eric E. Swayze
; APPLICANT: Lex M. Cowert
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: ISPH-0650
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: PCT/US01/02939
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 09/496,694
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/286,407
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 09/286,407
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 09/163,162
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 249
; SEQ ID NO 10
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (109)...(531)
US-10-181-316-10
Alignment Scores:
Pred. No.: 7,51e-79 Length: 955
Score: 660.00 Matches: 118
Percent Similarity: 92.86% Conservative: 12
Best Local Similarity: 84.29% Mismatches: 10
Query Match: 84.83% Indels: 0
DB: 13 Gaps: 0
US-09-690-825-34 (1-142) x US-10-181-316-10 (1-955)
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 109 ATGGGAGCTCCGGCGCTGCCAGATCTGGCAGCTGTACCTCAAGAACTACCGCATGCC 169
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 169 ACCTTCAGAACTGGCCCTCTCTGGAGGACTCGGCTGCACCCAGAGCGAATGGCGGAG 228
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db 229 GTGGCTTCATCCACTGCCCTACCGAGAACGAGCCCTGATTTGGCCAGTGTTTTCTGC 288
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80
Db 289 TTTAAGGAATTGGAGGCTGGAAACCCGATGACACCCGATAGAGAGCATGAAGACAC 348
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100
Db 349 TCCCTCGGCTGGCCCTCTCTACTGTCAAGAGAGCATGGAAAGAACTAACCGTCAGTGAA 408
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
Db 409 TTCTTGAAACTGGACAGACAGAGAGCCAGACAAATTTGAAAGAGAGCCACACACAG 468
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAla 140
Db 469 CAAAAAGAGTTTGAGAGAGACTGCAAGAGACTACCCGTCAGTCAATTGAGCAGCTGGCTGCC 528
RESULT 13
US-10-181-316-3
```

```
; Sequence 3, Application US/10181316
; Publication No. US20030211607A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Elizabeth J. Ackermann
; APPLICANT: Eric E. Swayze
; APPLICANT: Lex M. Cowert
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: ISPH-0650
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: PCT/US01/02939
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/496,694
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 09/286,407
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 09/163,162
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 249
; SEQ ID NO 3
; LENGTH: 14795
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2811)...(2921)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3174)...(3283)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5158)...(5275)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11955)...(12044)
US-10-181-316-3
Alignment Scores:
Pred. No.: 1.02e-39 Length: 14795
Score: 382.00 Matches: 73
Percent Similarity: 46.84% Conservative: 1
Best Local Similarity: 46.20% Mismatches: 0
Query Match: 49.10% Indels: 84
DB: 13 Gaps: 1
US-09-690-825-34 (1-142) x US-10-181-316-3 (1-14795)
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 2811 ATGGGTCCCGGAGCTTGCCCTGCTGGCAGCCCTTCTCTCAAGGACCAACCGCATCTCT 2870
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArg 37
Db 2871 ACATTCAAGAACTGGCCCTCTTGGAGGGCTGCCTGCACCCCGGAGCGGTGAGACTG 2930
QY 37 ----- 37
Db 2931 CCGGGCTCTGGGGTCCCCACGCCGCTTGCCTGTCTCCCTAGCGAGGCCACTGTGAC 2990
QY 37 ----- 37
Db 2991 TGGGCTCGGGGGTACAAAGCCGCTCTCCCTCCCGTCTGTCCCGAGCGGCACTGT 3050
QY 37 ----- 37
Db 3051 GGCTGGGCGCCCTTGGGTTCAGGGCGGGCTCCCTCCCTGCTGTTTGTCCCATCGAGGCCTT 3110
QY 37 ----- 37
Db 3111 TGTGGCTGGGCTCGGGGTTCCGGGGTTCGGCGGTGCACGCTCCACTCAGCAGCTGTGCTGCTCCTTG 3170
QY 38 ---MetAlaGluAlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGln 56
```

Db 3171 CAGATGCCGAGGCTGCTTCACTCCACTGCCCACTGAGAACGAGCAGACTTGGCCAG 3230  
Qy 57 CysPhePheCysPheCysGluLeuGluGlyTTPGluProAspAspProlle 74  
Db 3231 TGTTCCTCTCTCAAGGAGCTGGAAGGCTGGAGGCTGGAGCAGATGAGCAGCCCATG 3284  
RESULT 14  
US-09-954-456-973  
; Sequence 973, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; PRIOR FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 973  
; LENGTH: 14796  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-973  
Alignment Scores:  
Pred. No.: 1,02e-39 Length: 14796  
Score: 382.00 Matches: 73  
Percent Similarity: 46.84% Conservative: 1  
Best Local Similarity: 46.20% Mismatches: 0  
Query Match: 49.10% Indels: 84  
DB: 9 Gaps: 1  
US-09-690-825-34 (1-142) x US-09-954-456-973 (1-14796)  
Qy 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 2811 ATGGGTGCCCGGAGCTTCCCTGCTGGAGGCTTCTCAAGGACCCACCGCATCTCT 2870  
Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArg 37  
Db 2871 ACATTCAAGAACTGGCCCTTCTTGAGGGCTGCGCTGCACCCCGGAGCGGTGAGACTG 2930  
Qy 37 ----- 37  
Db 2931 CCGGCTCTCTGGGTGCCCGCCACGCGCCCTTGCCCTTGCCCTAGCGAGCCACTGTGAC 2990  
Qy 37 ----- 37  
Db 2991 TGGGCTCGGGGTACAAGCGCCCTCCCTCCCGTCTCTGTCGCCAGCGAGCCACTGT 3050  
Qy 37 ----- 37

Db 3051 GGCTGGGCCCCCTTGGGTCCAGGCGCGCTCCCTCCCTGCTTGTCCCATCGAGGCTT 3110  
Qy 37 ----- 37  
Db 3111 TGTGGCTGGGCTCGGGGTCCGGGCTGCCAGTCCACTCAGAGCTGTGCTGTCCCTTG 3170  
Qy 38 ---MetalAlaGluAlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGln 56  
Db 3171 CAGATGCCGAGGCTGGCTTCATCCACTGCCCCACTGAGAACGAGCAGACTTGGCCAG 3230  
Qy 57 CysPhePheCysPheLysGluLeuGluGlyTTPGluProAspAspProlle 74  
Db 3231 TGTTCCTCTCTCAAGGAGCTGGAAGGCTGGAGGCTGGAGCAGATGAGCAGCCCATG 3284  
RESULT 15  
US-09-954-456-1636  
; Sequence 1636, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; PRIOR FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1636  
; LENGTH: 14796  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-1636  
Alignment Scores:  
Pred. No.: 1,02e-39 Length: 14796  
Score: 382.00 Matches: 73  
Percent Similarity: 46.84% Conservative: 1  
Best Local Similarity: 46.20% Mismatches: 0  
Query Match: 49.10% Indels: 84  
DB: 9 Gaps: 1  
US-09-690-825-34 (1-142) x US-09-954-456-1636 (1-14796)  
Qy 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 2811 ATGGGTGCCCGGAGCTTCCCTGCTGGAGGCTTCTCAAGGACCCACCGCATCTCT 2870  
Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArg 37  
Db 2871 ACATTCAAGAACTGGCCCTTCTTGAGGGCTGCGCTGCACCCCGGAGCGGTGAGACTG 2930  
Qy 37 ----- 37

Db	2931	CCCGGCTCCTGGGGTCCCCACGCGCCTTGCCCTGTCCCTAGCGAGCCACTGTGAC	2990
QY	37	-----	37
Db	2991	TGGGCTCGGGGTACAGCGCCCTCCCTCCCGTCTGTCCCGAGGCGCACTGT	3050
QY	37	-----	37
Db	3051	GGCTGGGCCCCTTGGGTCCAGCGCGCCTCCCTCCCTGTTCCTCCCATCGAGGCCTT	3110
QY	37	-----	37
Db	3111	TGTGGTGGGCTCGGGGTTCGGGCTGCCACTCAGAGCTGTGTCCCTTG	3170
QY	38	---MetAlaGluAlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGln	56
Db	3171	CAGATGGCGGAGGCTGGCTTCATCCACTGCCCACTGAGACGAGCCAGACTTGGCCCGAG	3230
QY	57	CysPhePheCysPheLysGluLeuGlyTrpGluProAspAspProIle	74
Db	3231	TGTTTCTTCTGCTTCAGGAGCTGGAGGCTGGGAGCCAGATGACGCCCATG	3294

Search completed: August 16, 2004, 01:27:36  
Job time : 518 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 14:21:52 ; Search time 16 Seconds

(without alignments)  
853.700 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 142

Sequence: 1 MGAPLPPAWQFLLKDHRS.....EFETAKKVRRAIEQLAAMD 142

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

PIR 78:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	5.6	706	2 S74610	hypothetical prote
2	8	5.6	875	2 T27495	hypothetical prote
3	8	5.6	1822	2 S33441	EF protein - Strep
4	7	4.9	130	2 T15642	hypothetical prote
5	7	4.9	140	2 A57544	BM88 antigen - pig
6	7	4.9	152	2 AD1071	conserved hypotet
7	7	4.9	179	2 AD1597	20S proteasome bet
8	7	4.9	179	2 AF1234	20S proteasome bet
9	7	4.9	199	2 H71927	cag island protein
10	7	4.9	199	2 F64585	cag pathogenicity
11	7	4.9	228	2 D75048	aspartate racemase
12	7	4.9	228	2 G71112	probable aspartate
13	7	4.9	248	2 T13126	protein gp40 - pha
14	7	4.9	253	2 A96028	probable protocate
15	7	4.9	271	2 G69462	conserved hypotet
16	7	4.9	343	2 T40306	hypothetical prote
17	7	4.9	374	2 G88955	protein K04F1.6 [i
18	7	4.9	379	2 E75364	conserved hypotet
19	7	4.9	393	2 C83387	iron-sulfur cofact
20	7	4.9	427	2 T46148	hypothetical prote
21	7	4.9	442	2 A45529	heat shock protein
22	7	4.9	525	2 F97305	spore germination
23	7	4.9	662	2 T16845	hypothetical prote
24	7	4.9	698	2 AC0016	4-alpha-glucanotra
25	7	4.9	744	2 S64292	hypothetical prote
26	7	4.9	885	2 S76357	penicillin-binding
27	7	4.9	932	2 T04433	hypothetical prote
28	7	4.9	937	2 T43523	cut17 protein - fi
29	7	4.9	1098	2 T28816	hypothetical prote

30	7	4.9	2010	2 B71616	phosphatase (acid
31	7	4.9	2116	2 A26655	myosin heavy chain
32	7	4.9	2212	2 T28157	erythrocyte membra
33	7	4.9	2285	2 T12796	probable transglyc
34	7	4.9	2647	2 T28161	hypothetical prote
35	6	4.2	35	2 B95898	hypothetical prote
36	6	4.2	52	2 S77766	probable ornithine
37	6	4.2	62	2 S49522	homeotic protein e
38	6	4.2	62	2 S49558	hypothetical prote
39	6	4.2	68	2 H98098	hypothetical prote
40	6	4.2	73	2 AB1413	hypothetical prote
41	6	4.2	73	2 AI1788	hypothetical prote
42	6	4.2	75	2 S75054	hypothetical prote
43	6	4.2	80	2 AH2420	hypothetical prote
44	6	4.2	82	1 VCBPRK	coat protein B pre
45	6	4.2	87	2 B36907	pyocin S1 SII subu
46	6	4.2	87	2 D36907	pyocin S2 immunity
47	6	4.2	98	2 H96585	hypothetical prote
48	6	4.2	100	2 B71567	hypothetical prote
49	6	4.2	103	2 F71137	hypothetical prote
50	6	4.2	107	2 S53575	probable membrane
51	6	4.2	108	2 A71353	probable anti-sigm
52	6	4.2	108	2 T12829	probable Spbeta ph
53	6	4.2	108	2 F71158	hypothetical prote
54	6	4.2	115	2 T30042	hypothetical prote
55	6	4.2	116	2 C72265	hypothetical prote
56	6	4.2	121	2 B96847	hypothetical prote
57	6	4.2	122	2 AC2545	MobB protein limpo
58	6	4.2	131	2 D69145	conserved hypotet
59	6	4.2	133	2 AB3465	hypothetical prote
60	6	4.2	134	2 T35557	ribosomal protein
61	6	4.2	134	2 B75037	hypothetical prote
62	6	4.2	135	2 C71126	hypothetical prote
63	6	4.2	136	1 H64306	conserved hypotet
64	6	4.2	141	2 B64670	ribosomal protein
65	6	4.2	142	2 C72430	hypothetical prote
66	6	4.2	142	2 AF1517	hypothetical prote
67	6	4.2	145	2 B56150	probable general s
68	6	4.2	145	2 AB1277	transcription regu
69	6	4.2	149	2 H71205	hypothetical prote
70	6	4.2	150	2 AB1285	hypothetical prote
71	6	4.2	151	2 S43727	hypothetical prote
72	6	4.2	152	2 A87707	conserved hypotet
73	6	4.2	156	1 R3TW7	ribosomal protein
74	6	4.2	156	2 D75536	ribosomal protein
75	6	4.2	158	2 D82044	rRNA methylase, Sp
76	6	4.2	163	2 A72530	probable isosquino
77	6	4.2	168	2 A41547	ubiquitin-conjugat
78	6	4.2	170	2 B71864	hypothetical prote
79	6	4.2	172	2 E70057	hypothetical prote
80	6	4.2	172	2 E83886	hypothetical prote
81	6	4.2	172	2 AI2268	hypothetical prote
82	6	4.2	173	2 A90397	hypothetical prote
83	6	4.2	174	2 AI0013	heat shock protein
84	6	4.2	176	1 FWEC	inorganic diphosph
85	6	4.2	176	2 D91279	inorganic pyrophos
86	6	4.2	176	2 D86120	inorganic pyrophos
87	6	4.2	177	2 C87308	hypothetical prote
88	6	4.2	177	2 D83015	heat shock protein
89	6	4.2	178	2 S64357	hypothetical prote
90	6	4.2	182	2 F64569	hypothetical prote
91	6	4.2	188	2 T39211	hypothetical prote
92	6	4.2	196	2 PN0478	vancomycin resista
93	6	4.2	198	2 S50710	hypothetical prote
94	6	4.2	201	2 A31484	tropoin I, fast s
95	6	4.2	201	2 F70468	conserved hypotet
96	6	4.2	202	2 C87353	hypothetical prote
97	6	4.2	205	2 F83261	hypothetical prote
98	6	4.2	205	2 B97302	transcription regu
99	6	4.2	207	1 D64601	phosphoserine phos
100	6	4.2	207	2 B83523	hypothetical prote

## ALIGNMENTS

## RESULT 1

S74610  
 hypothetical protein slr1103 - *Synechocystis* sp. (strain PCC 6803)  
 C:Species: *Synechocystis* sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 17-Feb-2003  
 R:Accession: S74610  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* s.

A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S74610  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-706 <XAN>  
 A:Cross-references: EMBL:D90900; GB:AB001339; NID:gl651768; PIDN:BAA16762.1; PID:gl65183  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 C:Superfamily: signal transduction protein with FHA, Fas, GDEF, EAL domains

Query Match 5.6%; Score 8; DB 2; Length 706;  
 Best Local Similarity 100.0%; Pred. No. 5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 RRAIEQLA 139  
 DB 585 RRAIEQLA 592

## RESULT 2

T27495  
 hypothetical protein ZC15.1 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T27495  
 R: Basham, V.  
 submitted to the EMBL Data Library, March 1997

A:Reference number: Z20377  
 A:Accession: T27495  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-875 <WIL>  
 A:Cross-references: EMBL:Z93396; NID:el188462; PIDN:CAB07709.1; GSPDB:GN000023; CESP:ZC15  
 A:Experimental source: clone ZC15  
 C:Genetics:  
 A:Gene: CESP:ZC15.1  
 A:Map position: 5  
 A:Introns: 345/3; 380/3; 426/3; 669/2

Query Match 5.6%; Score 8; DB 2; Length 875;  
 Best Local Similarity 100.0%; Pred. No. 6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GAPTLPQA 9  
 DB 433 GAPTLPQA 440

## RESULT 3

S33441  
 EF protein - *Streptococcus suis*  
 C:Species: *Streptococcus suis*  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 15-Oct-1999  
 R:Smith, H.E.; Reek, F.H.; Veicht, U.; Gielkens, A.L.J.; Smits, M.A.  
 submitted to the EMBL Data Library, May 1993  
 A:Description: Repeats in an extracellular protein of weak-pathogenic strains are absent  
 A:Reference number: S33441  
 A:Accession: S33441

A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1822 <SMI>  
 A:Cross-references: EMBL:X71980; NID:G298031; PIDN:CAA50714.1; PID:G298032

Query Match 5.6%; Score 8; DB 2; Length 1822;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 AKNKIAKE 116  
 DB 1057 AKNKIAKE 1064

## RESULT 4

T15642  
 hypothetical protein C26F1.4 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 24-Nov-1999  
 C:Accession: T15642  
 R:Geisel, C.; Stellyes, L.; Bradshaw, H.  
 submitted to the EMBL Data Library, March 1996  
 A:Description: The sequence of *C. elegans* cosmid C26F1.  
 A:Reference number: Z18381  
 A:Accession: T15642  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-130 <GEI>  
 A:Cross-references: EMBL:U53148; NID:gl255375; PID:gl255381; PIDN:AAB37076.1; GSPDB:GN000001  
 A:Experimental source: strain Bristol N2; clone C26F1  
 C:Genetics:  
 A:Gene: CESP:C26F1.4  
 A:Map position: 5  
 A:Introns: 27/3; 71/1  
 C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 4.9%; Score 7; DB 2; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 BELTILGE 100  
 DB 47 BELTILGE 53

## RESULT 5

A57544  
 BM88 antigen - pig  
 C:Species: *Sus scrofa domestica* (domestic pig)  
 C:Date: 08-Feb-1996 #sequence\_revision 08-Feb-1996 #text\_change 05-Nov-1999  
 C:Accession: A57544; S49384  
 R:Mamalak, A.; Boutou, E.; Hurel, C.; Patsavoudi, E.; Tzartos, S.; Matsas, R.  
 J. Biol. Chem. 270, 14201-14208, 1995  
 A:Title: The BM88 antigen, a novel neuron-specific molecule, enhances the differentiation  
 A:Reference number: A57544; MUID:95294030; PMID:7775480

A:Accession: A57544  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-140 <MAM>  
 A:Cross-references: EMBL:X82027; NID:G557672; PIDN:CAA57549.1; PID:G557673  
 C:Superfamily: pig BM88 antigen  
 C:Keywords: phosphoprotein; transmembrane protein

Query Match 4.9%; Score 7; DB 2; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPQA 9  
 DB 50 APTLPQA 56

## RESULT 6

```

AD1071
conserved hypothetical protein SRY4902 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
A:Accession: AD1071
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AD1071
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-152 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD03387.1; PID:gi6505656; GSPDB:GN00176
C:Genetics:
A:Gene: SRY4902

Query Match 4.9%; Score 7; DB 2; Length 152;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 TPERWAE 40
      |||||
Db 135-TPERWAE 141

RESULT 7
20S proteasome beta-type chain homolog clpQ [imported] - Listeria innocua (strain Clp11
AD1597
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AD1597
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1597
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-179 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96548.1; PID:gi6413790; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: clpQ

Query Match 4.9%; Score 7; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 TAKKVR 133
      |||||
Db 36 TAKKVR 42

RESULT 8
AF1234
20S proteasome beta-type chain homolog clpQ [imported] - Listeria monocytogenes (strain
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AF1234
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

```

```

ok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1234
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-179 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC99356.1; PID:gi6410694; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: clpQ

Query Match 4.9%; Score 7; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 TAKKVR 133
      |||||
Db 36 TAKKVR 42

```

```

RESULT 9
H71927
cag island protein - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: H71927
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Merberg, D.; Millis, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; J
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: H71927
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-199 <ARN>
A:Cross-references: GB:AE001481; GB:AE001439; MID:94155005; PIDN:AAD06046.1; PID:94155006
A:Experimental source: strain J99
C:Genetics:
A:Gene: orf12

Query Match 4.9%; Score 7; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 ELTLGEF 101
      |||||
Db 98 ELTLGEF 104

```

```

RESULT 10
F64585
cag pathogenicity island protein cag6 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: F64585
R:Tomb, J.F.; White, O.; Kervavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: F64585
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-199 <TOM>
A:Cross-references: GB:AE000566; GB:AE000511; MID:92313628; PIDN:AAD07593.1; PID:92313640

```

```

Query Match 4.9%; Score 7; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

submitted to the EMBL Data Library, May 1998  
 A:Reference number: Z17603  
 A:Accession: T13126  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-248 <HEN>  
 A:Cross-references: EMBL:AF064539; NID:g3192683; PID:g3192714; PIDN:AAC19067.1  
 C:Genetics:  
 C:Superfamily: antitermination protein Q

Query Match 4.9%; Score 7; DB 2; Length 248;  
 Best Local Similarity 100.0%; Pred. No. 22; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0

QY 132 RRAIEQL 138  
 |||||  
 Db 91 RRAIEQL 97

RESULT 14  
 A:96028  
 probable protocatechuate 3,4-dioxygenase (EC 1.13.11.3) beta chain [imported] - Sinorhizobium meliloti  
 C:Species: Sinorhizobium meliloti  
 C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 20-Jan-2003  
 C:Accession: A96028  
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernandez, P.; Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A:Title: The complete sequence of the 1,683-Kb pSymB megaplasmid from the N2-fixing endosymbiont of the legume symbiont Sinorhizobium meliloti.  
 A:Reference number: A95842; MUID:21396508; PMID:11481431  
 A:Accession: A96028  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-253 <KUR>  
 A:Cross-references: GB:AL591985; PIDN:CAC49889.1; PID:g15141377; GSPDB:GN00167  
 A:Experimental source: strain 1021, megaplasmid pSymB  
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T. Science 293, 668-672, 2001  
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, H.; Neubert, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.R.; Wong, K.; Yeh, K.C.  
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: pcaH; SMD20577  
 A:Genome: plasmid  
 C:Superfamily: protocatechuate 3,4-dioxygenase beta chain  
 C:Keywords: oxidoreductase

Query Match 4.9%; Score 7; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 23; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0

QY 132 RRAIEQL 138  
 |||||  
 Db 209 RRAIEQL 215

RESULT 15  
 G69462  
 conserved hypothetical protein AF1704 - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 29-Sep-1999  
 C:Accession: G69462  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.; Nature 390, 364-370, 1997  
 A:Authors: Uitterlinden, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.P.; Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo-

QY 95 ETLTGEF 101  
 |||||  
 Db 98 ETLTGEF 104

RESULT 11  
 D75048  
 aspartate racemase PAB0912 - Pyrococcus abyssi (strain Orsay)  
 C:Species: Pyrococcus abyssi  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: D75048  
 R:Anonymous, Genoscope  
 submitted to the EMBL Data Library, July 1999  
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure  
 A:Reference number: A75001  
 A:Accession: D75048  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-228 <KAW>  
 A:Cross-references: GB:A0246287; GB:AL096836; NID:G5458657; PIDN:CAB50281.1; PID:G545879  
 A:Experimental source: strain Orsay  
 C:Genetics:  
 A:Gene: racD-1; PAB0912  
 C:Superfamily: Escherichia coli hypothetical 25.2K (lysR-araE intergenic region)

Query Match 4.9%; Score 7; DB 2; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 21; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0

QY 125 BETAKKV 131  
 |||||  
 Db 105 BETAKKV 111

RESULT 12  
 G71112  
 probable aspartate racemase - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hino, Y.; Yamamoto, S.; Sekin, M.; Ohtoku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon  
 A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: G71112  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-228 <KAW>  
 A:Cross-references: GB:AP000003; NID:g3236130; PIDN:BAA29761.1; PID:g3257078  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C:Genetics:  
 A:Gene: PH0670  
 C:Superfamily: Escherichia coli hypothetical 25.2K (lysR-araE intergenic region)

Query Match 4.9%; Score 7; DB 2; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 21; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0

QY 125 BETAKKV 131  
 |||||  
 Db 105 BETAKKV 111

RESULT 13  
 T13126  
 protein gp40 - phage N15  
 C:Species: phage N15  
 C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 03-Nov-2003  
 C:Accession: T13126  
 R:Hendrix, R.W.; Ravin, V.K.; Casjens, S.R.; Ford, M.E.; Ravin, N.V.; Smirnov, I.K.

A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: G69462  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-271 <KLB>  
A:Cross-references: GB:AE000986; GB:AE000782; NID:G2689309; PIDN:AAB89546.1; PID:G264885  
C:Superfamily: hypothetical protein HP0152

Query Match 4.9%; Score 7; DB 2; Length 271;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 LDRERAK 110  
|||||  
Db 220 LDRERAK 226

RESULT 16  
T40306  
hypothetical protein SPBC36.11 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T40306  
R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Moestl, D.; Duesterh  
submitted to the EMBL Data Library, May 1998  
A:Reference number: Z11919  
A:Accession: T40306  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-343 <LYN>  
A:Cross-references: EMBL:AL023589; PIDN:CAA19059.1; GSPDB:GN00067; SPDB:SPBC36.11  
A:Experimental source: strain 972h-; cosmid c36  
C:Genetics:  
A:Gene: SPDB:SPBC36.11  
A:Map position: 2

Query Match 4.9%; Score 7; DB 2; Length 343;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 EEKHKHS 81  
|||||  
Db 330 EEKHKHS 336

RESULT 17  
G88955  
protein K04F1.6 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: G88955  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: G88955  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-374 <STO>  
A:Cross-references: GB:chr\_V; PIDN:AACT8165.1; PID:G3886012; GSPDB:GN00023; CESP:K04F1.6  
C:Genetics:  
A:Gene: K04F1.6  
A:Map position: 5

Query Match 4.9%; Score 7; DB 2; Length 374;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 QFEELTL 98  
|||||  
Db 236 QFEELTL 242

RESULT 18

E75364  
conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: E75364  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; I  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: E75364  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-379 <WHI>  
A:Cross-references: GB:AE002011; GB:AE000513; NID:G6459461; PIDN:AAF11251.1; PID:G6459461  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR1694  
A:Map position: 1  
C:Superfamily: probable methyltransferase b0967

Query Match 4.9%; Score 7; DB 2; Length 379;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 AIEQLAA 140  
|||||  
Db 262 AIEQLAA 268

RESULT 19

C83387  
iron-sulfur cofactor synthesis protein PA2062 [similarity] - Pseudomonas aeruginosa (stra  
N:Contains: L-cysteine sulfurtransferase (EC 2.8.1.-)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: C83387  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bri  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:2043737; PMID:10984043  
A:Accession: C83387  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-393 <STO>  
A:Cross-references: GB:AE004633; GB:AE004091; NID:G9948070; PIDN:AG05450.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA2062  
C:Superfamily: nitrogen fixation protein nifs  
C:Keywords: phosphoprotein; pyridoxal phosphate; sulfurtransferase  
F:206/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted  
F:327/Active site: Cys (cysteine persulfide intermediate) #status predicted

Query Match 4.9%; Score 7; DB 2; Length 393;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PAWQPFL 14  
|||||  
Db 381 PAWQPFL 387

RESULT 20

T46148  
hypothetical protein T3A5.90 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
 C:Accession: T46148  
 R:Bloeker, H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quetier, F.; Sala  
 submitted to the Protein Sequence Database, December 1999  
 A:Reference number: Z23024  
 A:Accession: T46148  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-427 <BLO>  
 A:Cross-references: ENBL:AL132979  
 A:Experimental source: cultivar Columbia; BAC clone T3A5  
 C:Genetics:  
 A:Map position: 3  
 A:Introns: 286/3; 329/3  
 A:Note: T3A5.90

Query Match 4.9%; Score 7; DB 2; Length 427;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 71 DDPIEEH 77  
 |||||  
 Db 339 DDPIEEH 345

RESULT 21  
 A45529  
 heat shock protein 86 - fluke (Schistosoma mansoni) (fragment)  
 C:Species: Schistosoma mansoni  
 C>Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 20-Aug-1999  
 C:Accession: A45529  
 R:Johnson, K.S.; Wells, K.; Bock, J.V.; Nene, V.; Taylor, D.W.; Cordingley, J.S.  
 Mol. Biochem. Parasitol. 36, 19-28, 1989  
 A:Title: The 86-kilodalton antigen from Schistosoma mansoni is a heat-shock protein homolog  
 A:Reference number: A45529; MUID:90043865; PMID:2509907  
 A:Accession: A45529  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-442 <JOH>  
 A:Cross-references: GB:J04017; NID:G161027; PIDN:AAA29899.1; PID:G161028  
 C:Superfamily: heat shock protein 90

Query Match 4.9%; Score 7; DB 2; Length 442;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 90 KKQPEEL 96  
 |||||  
 Db 269 KKQPEEL 275

RESULT 22  
 F97305  
 spore germination protein GerKA, membrane protein [imported] - Clostridium acetobutylicum  
 C:Species: Clostridium acetobutylicum  
 C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
 C:Accession: F97305  
 R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
 A:Reference number: A96900; MUID:21359325; PMID:21359325  
 A:Accession: F97305  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-525 <KUR>  
 A:Cross-references: GB:AE001437; PIDN:AK81233.1; PID:G15026379; GSPDB:GN00168  
 A:Experimental source: Clostridium acetobutylicum ATCC824  
 C:Genetics:  
 A:Gene: CAC3300  
 C:Superfamily: spore germination protein gerBA

Query Match 4.9%; Score 7; DB 2; Length 525;

Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 BELTIGE 100  
 |||||  
 Db 183 BELTIGE 189

RESULT 23  
 T16845  
 hypothetical protein T10F2.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T16845  
 R:Miller, N.  
 submitted to the EMBL Data Library, March 1995  
 A:Description: The sequence of C. elegans cosmid T10F2.  
 A:Reference number: Z18589  
 A:Accession: T16845  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-662 <MIL>  
 A:Cross-references: EMBL:U23412; NID:G727446; PID:G727449; PIDN:AAA64293.1; CESP:T10F2.3  
 A:Experimental source: strain Bristol N2  
 C:Genetics:  
 A:Gene: CESP:T10F2.3  
 A:Introns: 17/1; 51/1; 96/3; 197/3; 342/3; 375/2; 619/2; 645/3; 649/1

Query Match 4.9%; Score 7; DB 2; Length 662;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 108 RAKNKIA 114  
 |||||  
 Db 339 RAKNKIA 345

RESULT 24  
 AC0016  
 4-alpha-glucanotransferase (EC 2.4.1.25) [imported] - Yersinia pestis (strain CO92)  
 C:Species: Yersinia pestis  
 C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AC0016  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
 Geno-Tarrag, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.  
 Nature 413, 523-527, 2001  
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: AC0016  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-698 <KUR>  
 A:Cross-references: GB:AL590842; PIDN:CAC89898.1; PID:G15978231; GSPDB:GN00175  
 C:Genetics:  
 A:Gene: malQ  
 C:Superfamily: 4-alpha-glucanotransferase  
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 4.9%; Score 7; DB 2; Length 698;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 LDRERAK 110  
 |||||  
 Db 586 LDRERAK 592

RESULT 25  
 S64292  
 hypothetical protein YGR003w - yeast (Saccharomyces cerevisiae)  
 A:Alternate names: hypothetical protein G3843  
 C:Species: Saccharomyces cerevisiae

C>Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 19-Apr-2002  
C/Accession: S64292  
R/Hebling, U.; Hofmann, B.; Delius, H.  
submitted to the Protein Sequence Database, May 1996  
A/Reference number: S64003  
A/Accession: S64292  
A/Molecule type: DNA  
A/Residues: 1-744 <HEB>  
A/Cross-references: EMBL:Z72788; NID:gl322956; PID:e243365; PID:gl322957; GSPDB:GN00007;  
A/Experimental source: strain S288C  
C/Genetics:  
A/Gene: MIPS:YGR003w  
A/Cross-references: SGD:S0003235  
A/Map position: 7R

Query Match 4.9%; Score 7; DB 2; Length 744;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 BELTLGE 100  
|||||  
Db 587 BELTLGE 593

## RESULT 26

S76357  
penicillin-binding protein 1A - *Synechocystis* sp. (strain PCC 6803)  
N/Alternate names: protein sll0002  
C/Species: *Synechocystis* sp.  
A/Variety: PCC 6803  
C/Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 01-Mar-2002  
C/Accession: S76357  
R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
s.  
A/Reference number: S74322; MUID:97061201; PMID:8905231  
A/Accession: S76357  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-885 <KAN>  
A/Cross-references: EMBL:D64000; GB:AB001339; NID:gl001484; PID:BAAL0209.1; PID:dl01086  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C/Genetics:  
A/Gene: ponA  
A/Start codon: GTG  
C/Superfamily: penicillin-binding protein

Query Match 4.9%; Score 7; DB 2; Length 885;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 BELTLGE 100  
|||||  
Db 370 BELTLGE 376

## RESULT 27

T04443  
hypothetical protein T18B16.190 - *Arabidopsis thaliana*  
C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
C/Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 30-Apr-1999  
C/Accession: T04443  
R/Bavan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Bancroft, I.; Mewes, H.W.  
submitted to the Protein Sequence Database, April 1998  
A/Reference number: Z15359  
A/Accession: T04443  
A/Molecule type: DNA  
A/Residues: 1-932 <BEV>  
A/Cross-references: EMBL:AL021687  
A/Experimental source: cultivar Columbia; BAC clone T18B16  
C/Genetics:

A/Map position: 4  
A/Introns: 527/2  
A/Note: T18B16.190

Query Match 4.9%; Score 7; DB 2; Length 932;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 BELTLGE 100  
|||||  
Db 303 BELTLGE 309

## RESULT 28

T43523  
cut17 protein - fission yeast (*Schizosaccharomyces pombe*)  
C/Species: *Schizosaccharomyces pombe*  
C/Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 02-Jun-2000  
C/Accession: T43523; T41649; T41700  
R/Morishita, J.; Matsusaka, T.; Yanagida, M.  
submitted to the EMBL Data Library, August 1999  
A/Description: Fission yeast cut17 is required for chromosome segregation.  
A/Reference number: Z22536  
A/Accession: T43523  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-997 <MOR>  
A/Cross-references: EMBL:AB031034; PIDN:BAAB3415.1  
R/Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, August 1998  
A/Reference number: Z22007  
A/Accession: T41649  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-997 <HAR>  
A/Cross-references: EMBL:AL031323; PIDN:CAA20434.1; GSPDB:GN00068; SPDB:SPCC962.02c  
A/Experimental source: strain 972h-; cosmid c962  
R/Wedler, H.; Duesterhoeft, A.; Mbougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, October 1999  
A/Reference number: Z22010  
A/Accession: T41700  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 932-997 <WED>  
A/Cross-references: EMBL:AL121859; PIDN:CAB59376.1; GSPDB:GN00068; SPDB:SPCP31B10.10c  
C/Genetics:  
A/Gene: cut17; SPCC962.02c; SPDB:SPCP31B10.10c  
A/Map position: 3L  
A/Introns: 43/3

Query Match 4.9%; Score 7; DB 2; Length 997;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 WEPPDDP 73  
|||||  
Db 173 WEPPDDP 179

## RESULT 29

T28816  
hypothetical protein F07C3.4 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C/Accession: T28816  
R/Favella, A.; Gattung, S.  
submitted to the EMBL Data Library, March 1996  
A/Description: The sequence of *C. elegans* cosmid F07C3.  
A/Reference number: Z20528  
A/Accession: T28816  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1098 <FAV>





A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A:Reference number: A69580; MUID:198044033; PMID:9384377  
A:Accession: A69911  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-2285 <KUN>  
A:Cross-references: GB:Z99115; GB:AL009126; NID:g2634479; PIDN:CAB14053.1; PID:e1183582;  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yomI

Query Match 4.9%; Score 7; DB 2; Length 2285;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0;

Qy 132 RRRAIEQL 138  
| | | | |  
Db 184 RRRAIEQL 190

RESULT 34  
T28161  
hypothetical protein FCR3-varT11-1 - malaria parasite (Plasmodium falciparum) (fragments  
C:Species: Plasmodium falciparum  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000  
A:Accession: T28161  
R:R Hernandez-Rivas, R.; Mattei, D.; Sterkers, Y.; Peterson, D.S.; Wellem, T.E.; Scherf, F.  
Mol. Cell. Biol. 17, 604-611, 1997  
A:Title: Expressed var genes are found in Plasmodium falciparum subtelomeric regions.  
A:Reference number: Z20483; MUID:97154495; PMID:9001213  
A:Accession: T28161  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2847 <HER>  
A:Cross-references: EMBL:U67959; NID:g1794255; PID:g1809295; PIDN:AAC47438.1  
A:Experimental source: strain FCQ27/PNG  
C:Genetics:  
A:Introns: 2158/3  
A:Note: FCR3-varT11-1

Query Match 4.9%; Score 7; DB 2; Length 2647;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 119 NKKKEPE 125  
| | | | |  
Db 382 NKKKEPE 388

RESULT 35  
E95098  
hypothetical protein SP0853 [imported] - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001  
A:Accession: E95098  
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
son, J.D.; Mayday, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
son, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: E95098  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-35 <KUR>  
A:Cross-references: GB:AE005672; PIDN:AAK74982.1; PID:g14972326; GSPDB:GN00164; TIGR:SP4  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP0853

Query Match 4.2%; Score 6; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 44;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 VKKQFE 94  
| | | | |  
Db 20 VKKQFE 25

RESULT 36  
S77766  
probable ornithine carbamoyltransferase (EC 2.1.3.3) - Mycoplasma capricolium (fragment)  
N:Alternate names: protein MC419  
C:Species: Mycoplasma capricolium  
C:Date: 09-Oct-1997 #sequence\_revision 24-Oct-1997 #text\_change 07-Dec-1999  
A:Accession: S77766  
R:Bork, P.; Ouzounis, C.; Casari, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert, W.;  
Mol. Microbiol. 16, 955-967, 1995  
A:Title: Exploring the Mycoplasma capricolium genome: a minimal cell reveals its physiolo  
A:Reference number: S77739; MUID:96059641; PMID:7476192  
A:Accession: S77766  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-52 <BOR>  
A:Cross-references: EMBL:Z23311; NID:G514486; PIDN:CAA83833.1; PID:G530484  
A:Experimental source: ATCC 27343  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
C:Genetics:  
A:Genetic code: SGC3  
C:Function:  
A:Description: catalyzes the conversion of ornithine and carbamoyl phosphate to citrullin  
A:Pathway: amino acid metabolism; arginine biosynthesis  
A:Note: evolutionary related to aspartate carbamoyltransferase (EC 2.1.3.2)  
C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase  
C:Keywords: arginine biosynthesis; transferase  
F:1-52/Domain: aspartate/ornithine carbamoyltransferase homology (fragment) <ACT>

Query Match 4.2%; Score 6; DB 2; Length 52;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 IAKETN 118  
| | | | |  
Db 18 IAKETN 23

RESULT 37  
S65922  
homeotic protein engrailed homolog A - Nautilus pompilius (fragment)  
C:Species: Nautilus pompilius  
C:Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 17-Mar-1999  
A:Accession: S65922  
R:Wray, C.G.; Jacobs, D.K.; Kostriken, R.; Vogler, A.P.; Baker, R.; Desallie, R.  
FEBS Lett. 365, 71-74, 1995  
A:Title: Homologues of the engrailed gene from five molluscan classes.  
A:Reference number: S65916; MUID:95293128; PMID:7774719  
A:Accession: S65922  
A:Molecule type: DNA  
A:Residues: 1-62 <WRA>  
A:Cross-references: EMBL:U23431  
C:Superfamily: unassigned homeobox proteins; homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F:1-50/Domain: homeobox homology (fragment) <HOX>

Query Match 4.2%; Score 6; DB 2; Length 62;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 KKEFEE 126  
| | | | |  
Db 9 KKEFEE 14

RESULT 38

S49558  
 Hypothetical protein 4 - Clostridium perfringens  
 C:Species: Clostridium perfringens  
 C:Date: 29-Nov-1995 #sequence\_revision 01-Mar-1996 #text\_change 15-Oct-1999  
 C:Accession: S49558  
 R:Lyristis, M.; Bryant, A.E.; Sloan, J.; Awad, M.M.; Nisbet, I.T.; Stevens, D.L.; Rood, M.J. Microbiol. 12, 761-777, 1994  
 A:Title: Identification and molecular analysis of a locus that regulates extracellular  
 A:Reference number: S49552; MUID:94328928; PMID:8052128  
 A:Accession: S49558  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-62 <LYR>  
 A:Cross-references: ENBL:U04966; NID:g498837; PIDN:AAA58952.1; PID:g498844  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1994

Query Match 4.2%; Score 6; DB 2; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FLSVKK 91  
 DB 39 FLSVKK 44  
 |||||

RESULT 39  
 H98098  
 Hypothetical protein spr1818 [imported] - Streptococcus pneumoniae (strain R6)  
 C:Species: Streptococcus pneumoniae  
 C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001  
 C:Accession: H98098  
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E  
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
 Y, P.; Sun, P.M.; Winkler, M.E.  
 J. Bacteriol. 183, 5709-5717, 2001  
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.  
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
 A:Reference number: A97872; MUID:21429245; PMID:11544234  
 A:Accession: H98098  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-68 <KUR>  
 A:Cross-references: GB:AE007317; PIDN:AAL00621.1; PID:g15459505; GSPDB:GN00174  
 C:Genetics:  
 A:Gene: spr1818

Query Match 4.2%; Score 6; DB 2; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KKEPEE 126  
 DB 63 KKEPEE 68  
 |||||

RESULT 40  
 AB1413  
 Hypothetical protein lmo2707 [imported] - Listeria monocytogenes (strain EGD-e)  
 C:Species: Listeria monocytogenes  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AB1413  
 R:Glaser, P.; Prangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
 D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Madueno, E.; Maitournam, A.; Ma  
 ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
 A:Title: Comparative Genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AB1413  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-73 <GLA>

A:Cross-references: GB:NC\_003210; PIDN:CAD00920.1; PID:g16412207; GSPDB:GN00177  
 A:Experimental source: strain EGD-e  
 C:Genetics:  
 A:Gene: lmo2707

Query Match 4.2%; Score 6; DB 2; Length 73;  
 Best Local Similarity 100.0%; Pred. No. 84;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 KIARET 117  
 DB 51 KIARET 56  
 |||||

RESULT 41  
 A11788  
 Hypothetical protein lin2855 [imported] - Listeria innocua (strain Clip11262)  
 C:Species: Listeria innocua  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: A11788  
 R:Glaser, P.; Prangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,  
 D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
 ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
 A:Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: A11788  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-73 <GLA>  
 A:Cross-references: GB:AL592022; PIDN:CAC98081.1; PID:g16415390; GSPDB:GN00178  
 A:Experimental source: strain Clip11262  
 C:Genetics:  
 A:Gene: lin2855

Query Match 4.2%; Score 6; DB 2; Length 73;  
 Best Local Similarity 100.0%; Pred. No. 84;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 KIARET 117  
 DB 51 KIARET 56  
 |||||

RESULT 42  
 S75054  
 Hypothetical protein ssl127 - Synecchocystis sp. (strain PCC 6803)  
 C:Species: Synecchocystis sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
 C:Accession: S75054  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
 O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,  
 DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocystis  
 s.  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S75054  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-75 <KAN>  
 A:Cross-references: ENBL:D90910; GB:AB001339; NID:g1652956; PIDN:BAAL7916.1; PID:d1018664  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 C:Genetics:  
 A:Start codon: GTG

Query Match 4.2%; Score 6; DB 2; Length 75;  
 Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 BELTIG 99

Db 69 EELTLG 74  
|||||

## RESULT 43

AH2420  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AH2420  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, Nakazaki, M.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AH2420  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-80 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA076619.1; PID:g17134058; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: asr4920

Query Match 4.2%; Score 6; DB 2; Length 80;

Best Local Similarity 100.0%; Pred. No. 91;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 EQLAAM 141  
|||||

Db 50 EQLAAM 55  
|||||

## RESULT 44

VCBPIX  
coat protein B precursor - phase Iike  
N:Alternate names: gene VII protein  
C:Species: phase Iike  
A:Note: host Escherichia coli  
C>Date: 29-Jul-1981 #sequence\_revision 28-May-1986 #text\_change 28-Jul-2000  
R:Peeters, B.P.H.; Peters, R.M.; Schoenmakers, J.G.G.; Konings, R.N.H.  
J. Mol. Biol. 131, 27-39, 1985  
A:Title: Nucleotide sequence and genetic organization of the genome of the N-specific filamentous bacteriophage phiX174  
A:Reference number: A92912; MUID:85160831; PMID:3981635  
A:Accession: A92912  
A:Molecule type: DNA  
A:Residues: 1-82 <PEE>  
A:Cross-references: GB:K02750; NID:g14942; PIDN:CAA26072.1; PID:g14948  
R:Nakashima, Y.; Frangione, B.; Wiseman, R.L.; Konigsberg, W.H.  
J. Biol. Chem. 256, 5792-5797, 1981  
A:Title: Primary structure of the major coat protein of the filamentous bacteriophage phiX174  
A:Reference number: A92326; MUID:81215498; PMID:7240173  
A:Accession: A92326  
A:Molecule type: protein  
A:Residues: 30-82 <NAK>  
A:Note: coat protein B is the major coat protein of the virion  
C:Genetics:  
A:Gene: VIII

C:Superfamily: filamentous phage coat protein B

C:Keywords: capsid protein

F1-29/Domain: signal sequence #status predicted <SIG>

F30-82/Product: coat protein B #status experimental <CPB>

Query Match 4.2%; Score 6; DB 1; Length 82;

Best Local Similarity 100.0%; Pred. No. 93;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 AKNKIA 114  
|||||

Db 10 AKNKIA 15  
|||||

## RESULT 45

B36907  
Pyocin S1 S11 subunit - Pseudomonas aeruginosa  
C:Species: Pseudomonas aeruginosa  
C>Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Jun-2000  
C:Accession: B36907  
R:Sano, Y.; Matsui, H.; Kobayashi, M.; Kageyama, M.  
J. Bacteriol. 175, 2907-2916, 1993  
A:Title: Molecular structures and functions of pyocins S1 and S2 in Pseudomonas aeruginosa  
A:Reference number: A36907; MUID:93259934; PMID:8491711  
A:Accession: B36907  
A:Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-87 <SAN>  
A:Cross-references: GB:D12707; NID:G286179; PIDN:BAA02202.1; PID:G286181  
A:Note: sequence extracted from NCBI backbone (NCBIN:131814, NCBI:P131817)  
C:Superfamily: immunity protein E2  
C:Keywords: bacteriocin immunity

Query Match 4.2%; Score 6; DB 2; Length 87;

Best Local Similarity 100.0%; Pred. No. 98;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 TNNKKK 122  
|||||

Db 22 TNNKKK 27  
|||||

## RESULT 46

D36907  
Pyocin S2 immunity protein S21 chain - Pseudomonas aeruginosa  
C:Species: Pseudomonas aeruginosa  
C>Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 31-Dec-2000  
C:Accession: D36907; E83501  
R:Sano, Y.; Matsui, H.; Kobayashi, M.; Kageyama, M.  
J. Bacteriol. 175, 2907-2916, 1993  
A:Title: Molecular structures and functions of pyocins S1 and S2 in Pseudomonas aeruginosa  
A:Reference number: A36907; MUID:93259934; PMID:8491711  
A:Accession: D36907  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-87 <SAN>  
A:Cross-references: GB:D12708; NID:G286182; PIDN:BAA02204.1; PID:G286184  
A:Note: sequence extracted from NCBI backbone (NCBI:P131820)  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brj  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: B83501  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-87 <STO>  
A:Cross-references: GB:AE004545; GB:AE004091; NID:g9947070; PIDN:AAG04540.1; GSPDB:GN001;  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: imm2; PA1151  
C:Superfamily: immunity protein E2  
C:Keywords: bacteriocin immunity

Query Match 4.2%; Score 6; DB 2; Length 87;

Best Local Similarity 100.0%; Pred. No. 98;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 TNNKKK 122  
|||||

Db 22 TNNKKK 27  
|||||

## RESULT 47

H96585

hypothetical protein F20D21.23 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: H96585  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.N.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Chen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Klaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Lin, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, J.C.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: H96585  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-98 <STO>  
 A:Cross-references: GB:AB005173; NID:94585983; PIDN:AAD25619.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1

Query Match 4.2%; Score 6; DB 2; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 EEKKH 80  
 |||||  
 DB 27 EEKKH 32

RESULT 48  
 B71567  
 hypothetical protein CT031 - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C:Species: Chlamydia trachomatis  
 C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 18-Aug-2000  
 C:Accession: B71567  
 R:Stephens, R.S.; Kaiman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.  
 A:Reference number: A71570; MUID:99000809; PMID:9784136  
 A:Accession: B71567  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-100 <ARN>  
 A:Cross-references: GB:AE001277; GB:AE001273; NID:93328408; PIDN:AAC67621.1; PID:9332842  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: CT031  
 C:Superfamily: conserved hypothetical protein TC0300

Query Match 4.2%; Score 6; DB 2; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 AKNKIA 114  
 |||||  
 DB 30 AKNKIA 35

RESULT 49  
 F71137  
 hypothetical protein PH0866 - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000  
 C:Accession: F71137  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekiri, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon.  
 A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: F71137  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-103 <KAW>  
 A:Cross-references: GB:AF000003; NID:93236130; PIDN:BAA29960.1; PID:93257277  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank C:Genetics:  
 A:Gene: PH0866  
 C:Superfamily: Pyrococcus horikoshii hypothetical protein PH0866

Query Match 4.2%; Score 6; DB 2; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 KHSGC 84  
 |||||  
 DB 53 KHSGC 58

RESULT 50  
 S53575  
 probable membrane protein YIR023c-a - Yeast (Saccharomyces cerevisiae)  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 06-May-1995 #sequence\_revision 19-Oct-1995 #text\_change 05-Nov-1999  
 C:Accession: S53575  
 R:Rowley, K.  
 submitted to the EMBL Data Library, October 1994  
 A:Reference number: S48478  
 A:Accession: S53575  
 A:Molecule type: DNA  
 A:Residues: 1-107 <ROW>  
 A:Cross-references: EMBL:Z38061; GSPDB:GN00009; MIPS:YIR023c-a  
 C:Genetics:  
 A:Gene: MIPS:YIR023c-a  
 A:Map position: 9R  
 C:Superfamily: Saccharomyces probable membrane protein YIR023c-a  
 C:Keywords: transmembrane protein  
 F10-26/Domain: transmembrane #status predicted <TW1>  
 F180-96/Domain: transmembrane #status predicted <TW2>

Query Match 4.2%; Score 6; DB 2; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 ELTIGE 100  
 |||||  
 DB 92 ELTIGE 97

RESULT 51  
 A71353  
 probable anti-sigma F factor antagonist (spoIIAA-1) - syphilis spirochete  
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
 C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
 C:Accession: A71353  
 R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwinrson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDorthey, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
 Science 281, 375-388, 1998  
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
 A:Reference number: A71250; MUID:98332770; PMID:9665876  
 A:Accession: A71353  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-108 <COL>  
 A:Cross-references: GB:AE001203; GB:AE000520; NID:93322476; PIDN:ARC65207.1; PID:93322487  
 A:Experimental source: strain Nichols  
 C:Genetics:  
 A:Gene: TP0220  
 C:Superfamily: sporulation protein stage II

Query Match 4.2%; Score 6; DB 2; Length 108;

Best Local Similarity 100.0%; Pred. No. 1.2e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0;

QY 131 VRRRAIE 136  
|||||

DB 85 VRRRAIE 90  
|||||

RESULT 52  
T12829  
probable SpBETA phage repressor - Bacillus subtilis phage SPBc2  
N:Alternate names: transcription regulator phage-related yonR  
C:Species: Bacillus subtilis phage SPBc2  
C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: T12829; A69315  
R:Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Mael, C.; Karamata, D.  
Submitted to the EMBL Data Library, August 1997  
A:Description: The complete nucleotide sequence of the Bacillus subtilis SPbetac2 prophage  
A:Reference number: Z17583  
A:Accession: T12829  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-108 <LAZ>  
A:Cross-references: EMBL:AF020713; NID:G3025478; PID:G3025543; PIDN:AAC13038.1  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Broutillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
A.; Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoisappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
Akeuchi, M.; Tamakoshi, A.; Tanaka, A.; Tanaka, P.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: A69915  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-108 <KUN>  
A:Cross-references: GB:Z99115; GB:AL009126; NID:G2634478; PIDN:CAB14020.1; PID:G2634522  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yonR  
C:Superfamily: probable transcription repressor yonR

QY 121 KKEFEE 126  
|||||

DB 102 KKEFEE 107  
|||||

RESULT 53  
F71158  
hypothetical protein PH0468 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
C:Accession: F71158  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A:Reference number: A71000; MUID:98344137; PMID:9679194  
A:Accession: F71158  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 1-108 <KAW>  
A:Cross-references: GB:AP000002; NID:G3236129; PIDN:BAA29555.1; PID:G3256872  
A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH0468

Query Match 4.2%; Score 6; DB 2; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0;

QY 106 RERAKN 111  
|||||

DB 57 RERAKN 62  
|||||

RESULT 54  
T30042  
hypothetical protein F20A1.10 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000  
C:Accession: T30042  
R:Gattung, S.; Wu, X.  
Submitted to the EMBL Data Library, March 1996  
A:Description: The sequence of C. elegans cosmid F20A1.  
A:Reference number: Z20726  
A:Accession: T30042  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-115 <GAT>  
A:Cross-references: EMBL:U53150; PIDN:AAA96132.1; GSPDB:GN00023; CESP:F20A1.10  
A:Experimental source: strain Bristol N2; clone F20A1  
C:Genetics:  
A:Gene: CESP:F20A1.10  
A:Map position: 5  
A:Introns: 25/1  
C:Superfamily: Caenorhabditis elegans hypothetical protein F57E7.1

Query Match 4.2%; Score 6; DB 2; Length 115;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 KKXFE 125  
|||||

DB 102 KKXFE 107  
|||||

RESULT 55  
C72265  
hypothetical protein - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: C72265  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
C.M.

Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: C72265  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-116 <ARN>  
A:Cross-references: GB:AE001789; GB:AE000512; NID:G4981904; PIDN:AAD36426.1; PID:G498191  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TMI355

Query Match 4.2%; Score 6; DB 2; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0;

QY 131 VRRRAIE 136  
|||||

```

Db      111 VRRRAIE 116
|||||
RESULT 56
B96647
hypothetical protein F19K23.2 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: B96647
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
anssen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B96647
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-121 <STO>
A:Cross-references: GB:AE005173; NID:g2160134; PIDN:AAB60756.1; GSPDB:GN00141
C:Genetics:
A:Gene: F19K23.2
A:Map position: 1

Query Match      4.2%; Score 6; DB 2; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 PTLPPA 9
|||||
Db      56 PTLPPA 61

RESULT 57
AC2545
McbB protein [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120beta
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AC2545
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC2545
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-122 <KUR>
A:Cross-references: GB:AP003602; PIDN:BA877289.1; PID:g17134731; GSPDB:GN00181
C:Genetics:
A:Gene: alr7646
A:Genome: plasmid

Query Match      4.2%; Score 6; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      114 AKETNN 119
|||||
Db      81 AKETNN 86

RESULT 58
D69145
conserved hypothetical protein MTH352 - Methanobacterium thermoautotrophicum (strain Del
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1999
C:Accession: D69145
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; E
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiawani, N.;
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi
A:Reference number: A69000; MUID:98037514; PMID:9371463
C:Accession: D69145
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-131 <MTH>
A:Cross-references: GB:AE000820; GB:AE000666; NID:g2621405; PIDN:AAB84858.1; PID:g2621410
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH352
C:Superfamily: hypothetical protein MJ1618

Query Match      4.2%; Score 6; DB 2; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      87 LSVKKQ 92
|||||
Db      123 LSVKKQ 128

RESULT 59
AB3465
hypothetical protein BMEI1704 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AB3465
R:Deivechio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, N.
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensi
A:Reference number: AD3252; PMID:11756688
A:Accession: AB3465
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-133 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52885.1; PID:g17983729; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1704
A:Map position: 1

Query Match      4.2%; Score 6; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      38 MAEAGF 43
|||||
Db      65 MAEAGF 70

RESULT 60
T35557
ribosomal protein S11 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 12-Jun-2003
C:Accession: T35557
R:Saunders, D.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1998
A:Reference number: Z21583
A:Accession: T35557
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-134 <SAU>
A:Cross-references: EMBL:AL031317; PIDN:CAA20384.1; GSPDB:GN00070; SCOEDB:SC6G4.06
A:Experimental source: strain A3(2)

```

C:Genetics:  
A:Gene: rpsK; SCOE:DB:SC6G4.06  
C:Superfamily: ribosomal protein S11/S14

Query Match 4.2%; Score 6; DB 2; Length 134;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 128 AKKVR 133  
Db 10 AKKVR 15  
|||||

RESULT 61  
B75037  
Hypothetical protein PAB0847 - Pyrococcus abyssi (strain Orsay)  
C:Species: Pyrococcus abyssi  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: B75037  
R:anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru  
A:Reference number: A75001  
A:Accession: B75037  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-134 <KAW>  
A:Cross-references: GB:AJ248287; GB:AL096836; NID:G5458657; PIDN:CAB50191.1; PID:G545870  
A:Experimental source: strain Orsay  
C:Genetics:  
A:Gene: PAB0847

Query Match 4.2%; Score 6; DB 2; Length 134;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 64 LEGWEP 69  
Db 42 LEGWEP 47  
|||||

RESULT 62  
C71126  
Hypothetical protein PH0777 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
C:Accession: C71126  
R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A:Reference number: A71000; MUID:98344137; PMID:9679194  
A:Accession: C71126  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-135 <KAW>  
A:Cross-references: GB:AP000003; NID:G3236130; PIDN:BA29869.1; PID:G3257186  
A:Experimental source: strain OT3  
A>Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH0777

Query Match 4.2%; Score 6; DB 2; Length 135;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 64 LEGWEP 69  
Db 40 LEGWEP 45  
|||||

RESULT 63  
H64306  
conserved hypothetical protein MJ0056 - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: H64306  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, C.  
ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.  
A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: H64306  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-136 <BUL>  
A:Cross-references: GB:U67463; GB:L77117; NID:G1590846; PIDN:AAB98036.1; PID:G1590850; TJ  
C:Genetics:  
A:Map position: REV56166-55756  
A:Start codon: TTG  
C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0056

Query Match 4.2%; Score 6; DB 1; Length 136;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 LKLDRE 107  
Db 46 LKLDRE 51  
|||||

RESULT 64  
B64570  
ribosomal protein L11 - Helicobacter pylori  
C:Species: Helicobacter pylori  
A:Variety: strains J99, 26895  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 24-Sep-1999  
C:Accession: B64570; F71846  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: B64570  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-141 <TCM>  
A:Cross-references: GB:AE000626; GB:AE000511; NID:G2314360; PIDN:AAD08248.1; PID:G2314364  
A:Experimental source: strain 26895  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; J  
Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A:Reference number: A71800; MUID:99120557; PMID:9923682  
A:Accession: F71846  
A:Molecule type: DNA  
A:Residues: 1-141 <ARN>  
A:Cross-references: GB:AE001540; GB:AE001439; NID:G4155717; PIDN:AAD06703.1; PID:G4155724  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: rplK; HPI202  
C:Superfamily: Escherichia coli ribosomal protein L11

Query Match 4.2%; Score 6; DB 2; Length 141;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 110 KXKIAK 115  
Db 94 KXKIAK 99  
|||||

## RESULT 65

C72430  
hypothetical protein - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 28-Jul-2000  
C:Accession: C72430  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
Garrett, M.E.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.  
C.M.

Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: C72430  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-142 <ARN>  
A:Cross-references: GB:AE001689; GB:AE000512; NID:g4980483; PIDN:AAD35103.1; PID:g498049  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM0009  
C:Superfamily: Thermotoga maritima hypothetical protein TM1526

Query Match 4.2%; Score 6; DB 2; Length 142;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 VRRRAIE 136  
| | | | |  
Db 96 VRRRAIE 101

## RESULT 66

AF1517  
hypothetical protein lin0678 [imported] - Listeria innocua (strain Clip11262)  
C:Species: Listeria innocua  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AF1517  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A:Title: Comparative Genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AF1517  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-142 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC95910.1; PID:g1641130; GSPDB:GN00178  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: lin0678

Query Match 4.2%; Score 6; DB 2; Length 142;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 GEFLKL 104  
| | | | |  
Db 112 GEFLKL 117

## RESULT 67

B56150  
probable general secretion pathway protein g precursor - Escherichia coli (strain K-12)  
N:Alternate names: hopG protein  
C:Species: Escherichia coli  
C>Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 01-Mar-2002  
C:Accession: B56150; C65126  
R:Stojiljkovic, I.; Schoenherr, R.; Kusters, J.G.  
J. Bacteriol. 177, 1892-1895, 1995  
A:Title: Identification of the hopG gene, a component of Escherichia coli K-12 type II e

A:Reference number: A56150; MUID:95204361; PMID:7896718

A:Accession: B56150  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-145 <STO>  
A:Cross-references: GB:U20786; NID:g693704; PIDN:AAA69031.1; PID:g693706  
R:Battner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col  
A.; Rose, D.J.; Mau, B.; Sha, Y.  
Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: C65126  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-145 <BLAT>  
A:Cross-references: GB:AE000409; GB:U00096; NID:g1789718; PIDN:AAC76353.1; PID:g1789725;  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: hofG  
C:Superfamily: secretion protein xcpT

Query Match 4.2%; Score 6; DB 2; Length 145;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 APTLPP 8  
| | | | |  
Db 79 APTLPP 84

## RESULT 68

AB1277  
transcription regulator MarR family homolog lmo1618 [imported] - Listeria monocytogenes  
C:Species: Listeria monocytogenes  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AB1277  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A:Title: Comparative Genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AB1277  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-146 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC99696.1; PID:g16411054; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo1618

Query Match 4.2%; Score 6; DB 2; Length 146;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 ELTLGE 100  
| | | | |  
Db 49 ELTLGE 54

## RESULT 69

H71205  
hypothetical protein PH1914 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000  
C:Accession: H71205  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine  
N.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic ar  
A:Reference number: A71000; MUID:98344137; PMID:9679194



A:Accession: H71205  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-149 <XAW>  
A:Cross-references: GB:AP000007; NID:G3236134; PIDN:BAA31039.1; PID:G3258356  
A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH1914  
C:Superfamily: Escherichia coli ybQ protein

Query Match 4.2%; Score 6; DB 2; Length 149;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 KEFEET 127  
|||||  
DB 60 KEFEET 65

RESULT 70  
AB1265  
Hypothetical protein lmo1522 [imported] - Listeria monocytogenes (strain EGD-e)  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
C:Accession: AB1265  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AB1265  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-150 <GLA>  
A:Cross-references: GB:NC 003210; PIDN:CAC99600.1; PID:G16410951; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo1522  
C:Superfamily: conserved hypothetical protein HI0670

Query Match 4.2%; Score 6; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AEAGFI 44  
|||||  
DB 114 AEAGFI 119

RESULT 71  
S43727  
Hypothetical protein - thermophilic bacterium PS-3 (fragment)  
C:Species: thermophilic bacterium PS-3  
C:Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 18-Sep-1998  
C:Accession: S43727  
R:Fujiwara, Y.; Oka, M.; Hamamoto, T.; Sone, N.  
Biochim. Biophys. Acta 1144, 213-219, 1993  
A:Title: Cytochrome c-551 of the thermophilic bacterium PS3, DNA sequence and analysis  
A:Reference number: S43726; MUID:93379042; PMID:7916623  
A:Accession: S43727  
A:Molecule type: DNA  
A:Residues: 1-151 <FUJ>  
A:Cross-references: EMBL:X63125  
C:Superfamily: conserved hypothetical protein yltI

Query Match 4.2%; Score 6; DB 2; Length 151;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAI 135  
|||||  
DB 80 KVRRAI 85

RESULT 72  
AB7707  
Conserved hypothetical protein CC3691 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: AB7707  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonitskii, J.; Esmailova, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: AB7249; MUID:21173698; PMID:11259647  
A:Accession: AB7707  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <STO>  
A:Cross-references: GB:AE005673; NID:G13425453; PIDN:AAK25653.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC3691

Query Match 4.2%; Score 6; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 ERMAEA 41  
|||||  
DB 94 ERMAEA 99

RESULT 73  
R3TW7  
Ribosomal protein S7 - Thermus aquaticus  
C:Species: Thermus aquaticus  
C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 12-Jun-2003  
C:Accession: S10250  
R:Yakhnin, A.V.; Vorozheykina, D.P.; Matvienko, N.I.  
Nucleic Acids Res. 18, 3659, 1990  
A:Title: Nucleotide sequence of the Thermus thermophilus HB8 rps12 and rps7 genes coding  
A:Reference number: S10249; MUID:90301504; PMID:2362824  
A:Accession: S10250  
A:Molecule type: DNA  
A:Residues: 1-156 <YAK>  
A:Cross-references: EMBL:X52165; NID:G48273; PIDN:CAA36419.1; PID:G48275  
A:Experimental source: strain HB8  
A:Note: the source is designated as Thermus thermophilus  
R:Tsiboli, P.; Herfurth, E.; Choi, T.  
Eur. J. Biochem. 226, 169-177, 1994  
A:Title: Purification and characterization of the 30S ribosomal proteins from the bacteri  
A:Reference number: S51053; MUID:95045586; PMID:7957245  
A:Accession: S51058  
A:Molecule type: protein  
A:Residues: 2-50 <TSI>  
A:Note: the source is given as Thermus thermophilus  
C:Superfamily: ribosomal protein S7  
C:Keywords: protein biosynthesis; ribosome

Query Match 4.2%; Score 6; DB 1; Length 156;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 ERMAEA 41  
|||||  
DB 142 ERMAEA 147

RESULT 74  
D75536  
ribosomal protein S7 - Deinococcus radiodurans (strain R1)

C;Species: Deinococcus radiodurans  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 12-Jun-2003  
 C;Accession: D75536  
 R;White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 S.; Smith, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
 Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A;Reference number: A75250; MUID:20036896; PMID:10567266  
 A;Accession: D75536  
 A;Status: Preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-156 <HEI>  
 A;Cross-references: GB:AE001891; GB:AE000513; NID:g6457973; PIDN:AAF09886.1; PID:g645797  
 A;Experimental source: strain R1  
 C;Genetics:  
 A;Gene: DR0306  
 A;Map position: 1  
 C;Superfamily: ribosomal protein S7

Query Match 4.2%; Score 6; DB 2; Length 156;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 ERVAA 41  
 |||||  
 Db 142 ERVAA 147

## RESULT 75

D82044  
 rRNA methylase, SpoU family VC2695 [imported] - Vibrio cholerae (strain N16961 serogroup  
 C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C;Accession: D82044  
 R;Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B  
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000  
 A;Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: D82044  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-158 <HEI>  
 A;Cross-references: GB:AE004335; GB:AE003852; NID:g9657289; PIDN:AAF95836.1; GSPDB:GN001  
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C;Genetics:  
 A;Gene: VC2695  
 A;Map position: 1  
 C;Superfamily: Chlamydomydia pneumoniae rRNA methylase

Query Match 4.2%; Score 6; DB 2; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 KVVRA 134  
 |||||  
 Db 42 KVVRA 47

Search completed: August 11, 2004, 14:25:32  
 Job time : 18 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 14:18:12 ; Search time 13 Seconds  
(without alignments)

568.766 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 142

Sequence: 1 MGAPTLPPAQPFLLKDHRS.....EFETAKKVRRAIEQLAAMD 142

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	142	100.0	142	1	BIR5_HUMAN
2	41	28.9	140	1	BIR5_MOUSE
3	39	27.5	142	1	BIR5_RAT
4	7	4.9	140	1	BM38_PIG
5	7	4.9	179	1	HSLV_LISIN
6	7	4.9	179	1	HSLV_LISMO
7	7	4.9	697	1	SENT_CAEEL
8	7	4.9	744	1	CUL3_YEAST
9	7	4.9	997	1	BIR1_SCHPO
10	7	4.9	2116	1	MY32_DICDI
11	6	4.2	78	1	ACP_TREHY
12	6	4.2	82	1	COM5_BPIKE
13	6	4.2	87	1	IMM1_PSEAE
14	6	4.2	87	1	IMM2_PSEAE
15	6	4.2	100	1	Y031_CHLTR
16	6	4.2	108	1	YGRB_BACPF
17	6	4.2	134	1	RS11_SPAW
18	6	4.2	134	1	RS11_STRCO
19	6	4.2	136	1	Y056_METJA
20	6	4.2	137	1	RL16_LEPIN
21	6	4.2	141	1	RL11_HELPY
22	6	4.2	145	1	GSPG_ECOLI
23	6	4.2	148	1	MSCL_PSESM
24	6	4.2	150	1	DTD_LISMO
25	6	4.2	150	1	YCCN_BACP3
26	6	4.2	155	1	RS7_THETH
27	6	4.2	156	1	RS7_DEIRA
28	6	4.2	157	1	Y839_HELHP
29	6	4.2	168	1	UBC7_WHEAT
30	6	4.2	173	1	HSLV_SHEON
31	6	4.2	173	1	HSLV_YERPE
32	6	4.2	175	1	HSLV_PSEPK
33	6	4.2	175	1	HSLV_PSESM

#### ALIGNMENTS

RESULT 1

34	4.2	175	1	IPYR_ECOL6	Q8fag0	escherichia
35	4.2	175	1	IPYR_ECOLI	P17288	escherichia
36	4.2	176	1	HSIV_PSEAE	Q9huc6	pseudomonas
37	4.2	178	1	YGV5_YEAST	P53071	saccharomyc
38	4.2	198	1	XJW7_YEAST	P40893	saccharomyc
39	4.2	201	1	RNH2_FORGI	O51832	porphyromon
40	4.2	201	1	TRI_FONLE	P05547	porphyromon
41	4.2	208	1	ZFP_IRV6	P47732	chilio iride
42	4.2	214	1	CYB6_EUGGR	P31480	mycoplasma gra
43	4.2	216	1	DEF_MYCCE	P47352	mycoplasma
44	4.2	217	1	YGV3_YEAST	P48238	saccharomyc
45	4.2	229	1	YMS1_ARCFU	O28033	archaeoglob
46	4.2	249	1	YGB1_FUSNN	O8rie0	fusobacteri
47	4.2	250	1	PUR7_SYNPK	Q7ub99	synthetococc
48	4.2	255	1	UPPS_METTH	O26334	methanobact
49	4.2	257	1	LIVG_ARCFU	O28881	archaeoglob
50	4.2	264	1	SPRC_CABEL	P34714	caenorhabdi
51	4.2	264	1	TRY6_ANOGA	P35040	anopheles g
52	4.2	273	1	SPSE_PYRFU	Q8u4g1	pyrococcus
53	4.2	281	1	Y765_HAEIN	O57125	haemophilus
54	4.2	286	1	FPG_STRCO	Q9zba6	streptomyce
55	4.2	286	1	SGBU_HAEIN	P44990	haemophilus
56	4.2	289	1	NOLT_RHIFR	P33209	rhizobium f
57	4.2	292	1	CHIA_CUCSA	P17541	cucumis sat
58	4.2	293	1	ENGC_BACAA	Q81wh7	bacillus an
59	4.2	294	1	Y237_MYCPN	P75455	mycoplasma
60	4.2	298	1	Y680_HAEIN	Q57389	haemophilus
61	4.2	305	1	LXD2_PHOLE	Q66878	photobacter
62	4.2	312	1	OTCC_MYCCC	P59779	mycoplasma
63	4.2	312	1	PANE_LACLA	Q9cfy8	lactococcus
64	4.2	314	1	CPKA_PYRAB	Q9uwf8	pyrococcus
65	4.2	318	1	NIA_CHLVU	Q01170	chlorella v
66	4.2	319	1	K6PE_OCEIH	Q8epd6	oceanobacil
67	4.2	324	1	ODPE_BACST	P21874	bacillus st
68	4.2	326	1	K6PF_MYCPU	Q98pw8	mycoplasma
69	4.2	340	1	HUR_STRAU	Q00923	streptomyce
70	4.2	340	1	RECA_MYCCE	P47581	mycoplasma
71	4.2	342	1	VANB_ENTFA	Q06893	enterococcu
72	4.2	342	1	VU7_HSV6U	Q01353	human herpe
73	4.2	347	1	MURE_VIBCH	Q9k440	vibrio chol
74	4.2	348	1	SYFA_STRPN	Q97836	streptococc
75	4.2	350	1	EGSA_SULTO	P58460	sulfolobus
76	4.2	352	1	NDR4_HUMAN	Q9ulp0	homo sapien
77	4.2	355	1	Y198_RICPR	Q9zdw7	rickettsia
78	4.2	361	1	CB45_MOUSE	O61112	mus musculu
79	4.2	361	1	CB45_RAT	Q91283	rattus norv
80	4.2	362	1	CB45_HUMAN	Q9brk5	homo sapien
81	4.2	365	1	CXAR_HUMAN	P78310	homo sapien
82	4.2	365	1	LEU3_BACCO	P12010	bacillus co
83	4.2	366	1	YQCC_BACSU	P45938	bacillus su
84	4.2	368	1	TRMU_ECO57	Q8x735	escherichia
85	4.2	368	1	TRMU_ECOL6	Q8cx28	escherichia
86	4.2	368	1	TRMU_ECOLI	P25745	escherichia
87	4.2	372	1	B4G2_HUMAN	O60909	h beta-1.4-
88	4.2	386	1	ICEA_XENIA	P55865	xenopus lae
89	4.2	387	1	YF20_METJA	Q58915	methanococc
90	4.2	390	1	YL35_CABEL	P34426	caenorhabdi
91	4.2	392	1	PORA_THEMA	O05651	thermotoga
92	4.2	399	1	ACKA_WIGER	Q8d320	wiggleswort
93	4.2	400	1	PBP2_NEIFL	P16873	neisseria f
94	4.2	401	1	ASSV_EUCAP	Q8ka60	baccharomyc
95	4.2	405	1	IFS_YEAST	P38431	saccharomyc
96	4.2	421	1	FXJ1_MOUSE	Q61660	mus musculu
97	4.2	421	1	FXJ1_RAT	Q63247	rattus norv
98	4.2	424	1	Y963_PYRHO	O58701	pyrococcus
99	4.2	425	1	SVS_THETN	Q8rdj5	thermoanaer
100	4.2	430	1	SECV_BACHD	P38375	bacillus ha

```

BIRS_HUMAN
ID BIRS_HUMAN STANDARD; PRT; 142 AA.
AC C15392; Q9P2W8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Baculoviral IAP repeat-containing protein 5 (Apoptosis inhibitor
DE survivin) (Apoptosis inhibitor 4).
CN BIRC5 OR API4 OR IAP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A. (ISOFORM ALPHA).
RP MEDLINE=97398388; PubMed=9256286;
RA Ambrosini G.; Adida C.; Altieri D.C.;
RT "A novel anti-apoptosis gene, survivin, expressed in cancer and
RT lymphoma."
RL Nat. Med. 3:917-921(1997).
RN (2)
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RC TISSUE=Pancratic carcinoma;
RA Uren A.G.; Vaux D.L.;
RT "Mammalian inhibitor of apoptosis (IAP) homolog."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A. (ISOFORM BETA).
RA Kagevama H.; Islam A.; Takayasu H.; Nakagawara A.;
RT "An isoform of survivin (survivin-beta) which has 23 amino acids
RT insertion into the BIR domain."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RC TISSUE=Lung, and Muscle;
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.;
RA Raha S.S.; Loquellano N.A.; Peters G.J.; Abramson R.D.; Mullany S.J.;
RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;
RA Villalón D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
RA Fahey J.; Helton E.; Kettelman M.; Madan A.; Rodrigues S.; Sanchez A.;
RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.;
RA Butterfield Y.S.N.; Krzywinski M.I.; Skalska U.; Smalish D.E.;
RA Schnerch A.; Schein J.E.; Jones S.J.M.; Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (5)
RP FUNCTION
RX MEDLINE=99075336; PubMed=9859999;
RA Li F.; Ambrosini G.; Chu E.Y.; Plescia J.; Tognin S.; Marchisio P.C.;
RA Altieri D.C.;
RT "Control of apoptosis and mitotic spindle checkpoint by survivin."
RL Nature 396:580-584(1998).
RN (6)
RP X-RAY CRYSTALLOGRAPHY (2.71 ANGSTROMS) OF ISOFORM ALPHA.
RX MEDLINE=20403315; PubMed=10949039;
RA Chantrelat L.; Skoufias D.A.; Klemm J.P.; Jung B.; Dideberg O.;
RA Margolis R.L.;
RT "Crystal structure of human survivin reveals a bow tie-shaped dimer
RT with two unusual alpha-helical extensions."
RL Mol. Cell 6:183-185(2000).
RN (7)
RP X-RAY CRYSTALLOGRAPHY (2.58 ANGSTROMS) OF ISOFORM ALPHA.
RX MEDLINE=20336902; PubMed=10876248;
RA Verdecia M.A.; Huang H.; Dutil E.; Kaiser D.A.; Hunter T.; Noel J.P.;
RT "Structure of the human anti-apoptotic protein survivin reveals a
RT dimeric arrangement."
RL Nat. Struct. Biol. 7:602-608(2000).
RN (8)
RP PHOSPHORYLATION OF THR-34.
RX MEDLINE=20542086; PubMed=11069302;
RA O'Connor D.S.; Grossman D.; Plescia J.; Li F.; Zhang H.; Villa A.;
RA Tognin S.; Marchisio P.C.; Altieri D.C.;
RT "Regulation of apoptosis at cell division by p34cdc2 phosphorylation
RT of survivin."
RL Proc. Natl. Acad. Sci. U.S.A. 97:13103-13107(2000).
CC -!- FUNCTION: May play a role in neoplasia. May counteract a default
CC induction of apoptosis in G2/M phase. Interacts with tubulin.
CC Inhibitor of caspase-3 and caspase-7.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Alpha;
CC IsoId=O15392-1; Sequence=Displayed;
CC Name=Beta;
CC IsoId=O15392-2; Sequence=VSP_002454;
CC -!- TISSUE SPECIFICITY: Expressed only in fetal kidney and liver, and
CC to lesser extent, lung and brain. Abundantly expressed in
CC adenocarcinoma (lung, pancreas, colon, breast, and prostate) and
CC in high-grade lymphomas.
CC -!- SIMILARITY: Contains 1 BIR repeat.
CC -!- SIMILARITY: Belongs to the IAP family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U75285; AAC51660.1; -
CC EMBL; AF077350; AAD34226.1; -
CC EMBL; AB028869; BA93676.1; -
CC EMBL; BC008718; AA080718.1; -
CC EMBL; BC034148; AA034148.1; -
CC PDB; 1B31; 03-JAN-01.
CC PDB; 1F3H; 16-MAY-01.
CC Genew; HGNC:593; BIRC5.
CC MIM; 603352; -
CC GO; GO:0005876; C:spindle microtubule; TAS.
CC GO; GO:0008189; P:apoptosis inhibitor activity; TAS.
CC GO; GO:0006916; P:anti-apoptosis; TAS.
CC GO; GO:0000086; P:G2/M transition of mitotic cell cycle; TAS.
CC InterPro; IPR001370; BIR.
CC Pfam; PF00653; BIR; 1.
CC SMART; SM00238; BIR; 1.
CC PROSITE; PS01262; BIR_REPEAT_1; FALSE_NEG.
CC PROSITE; PS0143; BIR_REPEAT_2; 1.
CC Apoptosis; thiol protease inhibitor; Alternative splicing;
CC Metal-binding; Zinc; Phosphorylation; 3D-structure.
CC REPEAT 15 87
CC METAL 57 57 ZINC.
CC METAL 60 60 ZINC.
CC METAL 77 77 ZINC.
CC METAL 84 84 ZINC.
CC MOD_RES 34 34
CC VARSPLIC 74 74
CC
CC MUTAGEN 84 84
CC HELIX 8 13
CC HELIX 15 20
CC TURN 21 21
CC TURN 29 30

```

FT HELIX 35 40  
TURN 41 42  
FT STRAND 43 45  
FT STRAND 45 57  
FT TURN 55 57  
FT STRAND 58 60  
FT STRAND 63 64  
FT TURN 69 70  
FT TURN 73 80  
FT TURN 82 83  
FT TURN 85 87  
FT TURN 88 88  
FT TURN 93 95  
FT STRAND 97 97  
FT HELIX 98 139  
SQ SEQUENCE 142 AA; 16389 NW; 9E7CADCDF2822286 CRC64;  
  
Query Match 100.0%; Score 142; DB 1; Length 142;  
Best Local Similarity 100.0%; Pred. No. 7.2e-140;  
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MGAPTLPPAWQFLLKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTNEPDLAQCFPC 60  
Db 1 MGAPTLPPAWQFLLKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTNEPDLAQCFPC 60  
  
QY 61 FKELEGWEPDDPIIEHKHSGGCAFLSVKKQFELTIGFELKDRERAKNKIAKETNNK 120  
Db 61 FKELEGWEPDDPIIEHKHSGGCAFLSVKKQFELTIGFELKDRERAKNKIAKETNNK 120  
  
QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
Db 121 KKEFEETAKKVRRAIEQLAAMD 142  
  
RESULT 2  
BIRS MOUSE STANDARD; PRT: 140 AA.  
ID 070201; Q923F7; Q9WU53; Q9WU54;  
AC 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE Baculoviral IAP repeat-containing protein 5 (Apoptosis inhibitor  
DE survivin) (Apoptosis inhibitor 4) (TIAP).  
GN BIR5 OR API4 OR IAP4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RN SEQUENCE FROM N.A. (ISOFORM 1).  
RP Uren A.G., Vaux D.L.;  
RA "Mammalian inhibitor of apoptosis (IAP) homolog.";  
RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
RL [2]  
RN SEQUENCE FROM N.A. (ISOFORM 1).  
RP TISSUE=Embryo;  
RA Kobayashi K., Otaki M., Ogasawara T., Tokuhisa T.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RX MEDLINE=20129603; PubMed=1066622;  
RA Conway E.M., Pollefeys S., Cornilissen J., DeBaere I.,  
RA Steiner-Mosonyi M., Ong K., Baens M., Collen D., Schu A.C.;  
RT "Three differentially expressed survivin cDNA variants encode  
RT proteins with distinct antiapoptotic functions.";  
RL Blood 95:1435-1442(2000).  
RN [4]  
RN SEQUENCE FROM N.A. (ISOFORM 2).  
RX MEDLINE=22389257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Rahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [5]  
RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 7-118, AND PHOSPHORYLATION  
OF THR-34.  
RP MEDLINE=20403314; PubMed=10949038;  
RX Muchmore S.W., Chen J., Jakob C., Zakula D., Matayoshi E.D., Wu W.,  
RA Zhang H., Li F., Ng S.C., Altieri D.C.;  
RA "Crystal structure and mutagenic analysis of the  
RT inhibitor-of-apoptosis protein survivin";  
RL Mol. Cell 6:173-182(2000).  
CC -!- FUNCTION: May play a role in neoplasia. May counteract a default  
CC induction of apoptosis in G2/M phase. Interacts with tubulin.  
CC Inhibitor of caspase-3 and caspase-7 (By similarity).  
CC -!- SUBUNIT: Homodimer; zinc-dependent.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=1; Synonyms=Survivin 140;  
CC IsoId=O70201-1; Sequence=Displayed;  
CC Name=2; Synonyms=Survivin 121;  
CC IsoId=O70201-2; Sequence=VSP\_002457;  
CC Name=3; Synonyms=Survivin 40;  
CC IsoId=O70201-3; Sequence=VSP\_002455, VSP\_002456;  
CC -!- SIMILARITY: Contains 1 BIR repeat.  
CC -!- SIMILARITY: Belongs to the IAP family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF077343; AAD34225.1; -  
DR EMBL; AB013819; BAA28266.1; -  
DR EMBL; AF115517; AAD26199.1; -  
DR EMBL; AF115517; AAD26200.1; -  
DR EMBL; AF115517; AAD26201.1; -  
DR EMBL; BC004702; AAH04702.1; -  
DR PDB; 1M4M; 25-SEP-02.  
DR MGD; MGI-1203517; Birc5.  
DR GO; GO:0005737; Cytoplasm; IDA.  
DR GO; GO:0008189; F:apoptosis inhibitor activity; IMP.  
DR GO; GO:0006916; P:anti-apoptosis; IMP.  
DR InterPro; IPR001370; BIR.  
DR Pfam; PF00653; BIR; 1.  
DR SMART; SM00238; BIR; 1.  
DR PROSITE; PS01282; BIR\_REPEAT\_1; FALSE\_NEG.  
DR PROSITE; PS50143; BIR\_REPEAT\_2; 1.  
DR Apoptosis; Thiol protease inhibitor; Alternative splicing;  
KW Phosphorylation; Metal-binding; Zinc; 3D-structure.  
FT REPEAT 15 87  
FT MOD\_RES 34 34 PHOSPHORYLATION (BY CDC2).  
FT METAL 57 57 ZINC 1.  
FT METAL 60 60 ZINC 1.  
FT METAL 76 76 ZINC 2.  
FT METAL 77 77 ZINC 1.  
FT METAL 80 80 ZINC 2.

```

FT METAL 84 84 ZINC 1.
FT VARSPLIC 38 40 MAE -> RGA (in isoform 3).
FT VARSPLIC 41 140 /FTid=VSP_002455.
FT VARSPLIC 114 140 Missing (in isoform 3).
FT VARSPLIC 114 140 /FTid=VSP_002456.
FT VARSPLIC 114 140 AKETNNKKEEETAKTTTQRIQIAA -> VCMIEKND
FT VARSPLIC 114 140 (in isoform 2).
FT VARSPLIC 114 140 /FTid=VSP_002457.
FT VARSPLIC 114 140 /FTid=VSP_002458.
SQ SEQUENCE 140 AA; 16297 MW; 28F5ABF501A6D83C CRC64;

Query Match 28.9%; Score 41; DB 1; Length 140;
Best Local Similarity 100.0%; Pred. No. 3.9e-35;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 CACTPERVAEAGFTHCPTENEPDLAQCFFCFKELEGWEPDD 71
Db 31 CACTPERVAEAGFTHCPTENEPDLAQCFFCFKELEGWEPDD 71

RESULT 3
BIRS RAT
ID BIRS RAT STANDARD; PRT; 142 AA.
AC Q9JHY7,
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Baculoviral IAP repeat-containing protein 5 (Apoptosis inhibitor
DE survivin).
GN BIRC5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Chen D., Cao G., Chen J.;
RT Molecular cloning and characterization of rat survivin.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May play a role in neoplasia. May counteract a default
CC induction of apoptosis in G2/M phase. Interacts with tubulin.
CC Inhibitor of caspase-3 and caspase-7 (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 BIR repeat.
CC -!- SIMILARITY: Belongs to the IAP family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF276775; AAF82586.1; -
CC HSPSP; O15392; IEF31.
CC InterPro; IPR01370; BIR.
CC Pfam; PF00653; BIR; 1.
CC SMART; SM00238; BIR; 1.
CC PROSITE; PS01282; BIR_REPEAT_1; FALSE_NEG.
CC PROSITE; PS0143; BIR_REPEAT_2; 1.
CC Apoptosis; Thiol protease inhibitor; Metal-binding; Zinc;
KW Phosphorylation.
FT REPEAT 15 87 BIR.
FT METAL 57 57 ZINC (BY SIMILARITY).
FT METAL 60 60 ZINC (BY SIMILARITY).
FT METAL 77 77 ZINC (BY SIMILARITY).
FT METAL 84 84 ZINC (BY SIMILARITY).
FT MOD_RES 34 34 PHOSPHORYLATION (BY CDC2)
FT (BY SIMILARITY).
SQ SEQUENCE 142 AA; 16692 MW; 99BCFAE1584D0CAC CRC64;

```

```

Query Match 27.5%; Score 39; DB 1; Length 142;
Best Local Similarity 100.0%; Pred. No. 4.7e-33;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CTPERVAEAGFTHCPTENEPDLAQCFFCFKELEGWEPDD 71
Db 33 CTPERVAEAGFTHCPTENEPDLAQCFFCFKELEGWEPDD 71

RESULT 4
BM88 PIG
ID BM88 PIG STANDARD; PRT; 140 AA.
AC Q290Z6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE BM88 antigen.
GN BM88.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 7-17, AND CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=95294030; PubMed=7775480;
RA Mamalaki A., Boucou E., Hurel C., Patsavoudi E., Tzartos S.,
RA Matsas R.;
RT "The BM88 antigen, a novel neuron-specific molecule, enhances the
RT differentiation of mouse neuroblastoma cells.";
RL J. Biol. Chem. 270:14201-14208(1995).
CC -!- FUNCTION: Involved in neuroblastoma cell differentiation.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Type IV membrane protein. Associated with
CC the plasma membrane and the membranes of intracellular organelles
CC (Potential).
CC -!- TISSUE SPECIFICITY: Found in the spinal cord, cerebellum and
CC cerebrium, where it is localized in neurons.
CC -!- DEVELOPMENTAL STAGE: Appears during early brain development where
CC it increases with age to give high levels in the mature animal.
CC -!- SIMILARITY: BELONGS TO THE BM88 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X82027; CAA57549.1; -
CC PIR; A57544; A57544.
CC Antigen; Transmembrane.
KW DOMAIN 1 115
FT TRANSMEM 116 136 CYTOPLASMIC (POTENTIAL).
FT ANCHOR FOR TYPE IV MEMBRANE PROTEIN
FT (POTENTIAL).
FT DOMAIN 137 140 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 123 128 POLY-ALA.
SQ SEQUENCE 140 AA; 13968 MW; 1F4F6C2D2FEC6D6D CRC64;

Query Match 4.9%; Score 7; DB 1; Length 140;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPPA 9
Db 50 APTLPPA 56

RESULT 5
HSLV LISIN STANDARD; PRT; 179 AA.
ID HSLV LISIN
AC Q92C74;

```

```

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ATP-dependent protease hslv (EC 3.4.25.-).
GN HSLV OR CLPQ OR L1N1317.
OS Listeria innocua.
OX Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
CX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul P., Bloecher H., Brandt P., Chakraborty T.,
RA Baquero F., Berche P., Couve E., de Paruvar A., Dehoux P.,
RA Charbit A., Chetouani F., Duchaud E., Durant L., Dussurget O.,
RA Domann E., Dominguez-Bernal G., Garcia-del Portillo F., Garrido P.,
RA Entian K.-D., Fshih H., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Kurapkat G.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Nedjari H.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RL "Comparative genomics of Listeria species.";
Science 294:849-852(2001).
CC -!- FUNCTION: Protease subunit of a proteasome-like degradation
CC complex (By similarity).
CC -!- SUBUNIT: A double ring-shaped homohexameric of hslv is capped on
CC each side by a ring-shaped hslu homohexameric (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family T1B. Hslv subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AL596168; CAC96548.1; -.
CC F1R; AD1597; AD1597.
CC MEROPS; T01.007; -.
CC Listlist; L1N01317; -.
CC HAMAP; MF 00248; -.
CC InterPro; IPR001353; Peptidase_T1.
CC Pfam; PF00227; Proteasome; 1.
CC Hydrolase; Protease; Threonine protease; Complete proteome.
CC ACT SITE 6 BY SIMILARITY.
CC SQ SEQUENCE 179 AA; 19390 MW; D546A0476C5662B1 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 179;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 TAKKVR 133
Db 36 TAKKVR 42

RESULT 6
HSLV LISMO STANDARD; PRT; 179 AA.
AC Q8Y7J9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ATP-dependent protease hslv (EC 3.4.25.-).
GN HSLV OR CLPQ OR LMO1276.
OS Listeria monocytogenes.
OX Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
CX NCBI_TaxID=1639;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul P., Bloecher H., Brandt P., Chakraborty T.,
RA Baquero F., Berche P., Couve E., de Paruvar A., Dehoux P.,
RA Charbit A., Chetouani F., Duchaud E., Durant L., Dussurget O.,
RA Domann E., Dominguez-Bernal G., Garcia-del Portillo F., Garrido P.,
RA Entian K.-D., Fshih H., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Kurapkat G.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Nedjari H.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RL "Comparative genomics of Listeria species.";
Science 294:849-852(2001).
CC -!- FUNCTION: Protease subunit of a proteasome-like degradation
CC complex (By similarity).
CC -!- SUBUNIT: A double ring-shaped homohexameric of hslv is capped on
CC each side by a ring-shaped hslu homohexameric (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family T1B. Hslv subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AL591978; CAC99356.1; -.
CC F1R; AF1234; AF1234.
CC MEROPS; T01.007; -.
CC Listlist; LMO01278; -.
CC HAMAP; MF 00248; -.
CC InterPro; IPR001353; Peptidase_T1.
CC Pfam; PF00227; Proteasome; 1.
CC Hydrolase; Protease; Threonine protease; Complete proteome.
CC ACT SITE 6 BY SIMILARITY.
CC SQ SEQUENCE 179 AA; 19406 MW; 443CB9B915B1FCC8 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 179;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 TAKKVR 133
Db 36 TAKKVR 42

RESULT 7
SENT_CABEL STANDARD; PRT; 697 AA.
ID AC Q09353; Q8IU18;
DT 01-NOV-1997 (Rel. 35, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable sentrin-specific protease (EC 3.4.22.-).
GN T10F2.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Tsubaki A., Incue H., Takahashi K.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
CX Miller N.;
RN [1]

```





```

RX MEDLINE=99398681; PubMed=10468581;
RA Uren A.G., Beilharz T., O'Connell M.J., Bugg S.J., van Driel R.,
RA Vaux D.L., Lithow T.,
RT "Role for yeast inhibitor of apoptosis (IAP)-like proteins in cell
RT division.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10170-10175(1999).
RN [4]
RN CHARACTERIZATION.
RP MEDLINE=21850422; PubMed=11861551;
RA Rajagopalan S., Balasubramanian M.K.;
RA "Schizosaccharomyces pombe Birp, a nuclear protein that localizes to
RT kinetochores and the spindle midzone, is essential for chromosome
RT condensation and spindle elongation during mitosis.";
RL Genetics 160:445-456(2002).
RN [5]
RN FUNCTION.
RP MEDLINE=20035862; PubMed=10571085;
RA Rajagopalan S., Balasubramanian M.K.;
RA "S. pombe Birp: an inhibitor of apoptosis domain containing protein
RT is essential for chromosome segregation.";
RL FEBS Lett. 460:187-190(1999).
RN CC
RN CC -!- FUNCTION: Seems to act in the pleiotropic control of cell
RN CC division. Has a role in chromosome segregation by recruiting
RN CC condensin and ark1 kinase to appropriate sites as the cell
RN CC progresses through mitosis.
RN CC -!- SUBCELLULAR LOCATION: Nuclear. Interacts with the outer
RN CC centromeric regions of the chromosomes during interphase. After
RN CC chromatid separation moves to the middle of the spindle.
RN CC -!- SIMILARITY: Contains 2 BIR repeats.
RN CC
RN CC -----
RN CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RN CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN CC the European Bioinformatics Institute. There are no restrictions on its
RN CC use by non-profit institutions as long as its content is in no way
RN CC modified and this statement is not removed. Usage by and for commercial
RN CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
RN CC or send an email to license@isb-sib.ch).
RN CC
RN CC -----
RN CC EMBL; AB031034; BAA83415.1; -
RN CC EMBL; AL031323; CAA20434.1; -
RN CC EMBL; AL121859; CAB58376.1; -
RN CC PIR; T43523; T43523.
RN CC HSP; Q13490; 1QBH.
RN CC GenDB_SPombe; SPCC962.02c; -
RN CC InterPro; IPR001370; BIR.
RN CC Pfam; PF00653; BIR; 2.
RN CC SMART; SM00238; BIR; 2.
RN CC PROSITE; PS01282; BIR_REPEAT 1; FALSE_NEG.
RN CC PROSITE; PS0143; BIR_REPEAT 2; 2.
RN CC Cell division; Mitosis; Nuclear protein; Repeat.
KW REPEAT 25 39 BIR 1.
FT REPEAT 120 194 BIR 2.
FT DOMAIN 80 83 POLY-ASP.
FT DOMAIN 312 319 POLY-ASP.
FT DOMAIN 487 490 POLY-SER.
SQ SEQUENCE 997 AA; 112579 MW; 952A6BAFA5C489F4 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 997;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 WEPDDDP 73
Db 173 WEPDDDP 179
|||||
|||||

RESULT 10
MYS2 DICDI STANDARD; PRT; 2116 AA.
AC P08759;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)

```

```

DE Myosin II heavy chain, non muscle.
GN MHCA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=87092266; PubMed=3540939;
RA Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;
RT "Conserved protein domains in a myosin heavy chain gene from
RT Dictyostelium discoideum.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN [2]
RN PHOSPHORYLATION SITES, AND MUTAGENESIS.
RP STRAIN=AX2;
RC MEDLINE=90353583; PubMed=2387408;
RA Lueck-Vielmeier D., Schleicher M., Grabatin B., Wippler J.,
RA Gerisch G.;
RT "Replacement of threonine residues by serine and alanine in a
RT phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
RL FEBS Lett. 269:239-243(1990).
RN [3]
RN PHOSPHORYLATION SITES.
RP MEDLINE=88112226; PubMed=2828113;
RA Wagle G., Noegel A., Scheel J., Gerisch G.;
RT "Phosphorylation of threonine residues on cloned fragments of the
RT Dictyostelium myosin heavy chain.";
RL FEBS Lett. 227:71-75(1988).
RN [4]
RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RP MEDLINE=95345066; PubMed=7619795;
RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
RA Rayment I.;
RT "X-ray structures of the myosin motor domain of Dictyostelium
RT discoideum complexed with MgADP.Befx and MgADP.ALFA-.";
RL Biochemistry 34:8960-8972(1995).
RN [5]
RN X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RP MEDLINE=95345067; PubMed=7619796;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the
RT truncated head of Dictyostelium discoideum myosin to 2.7-A
RT resolution.";
RL Biochemistry 34:8973-8981(1995).
RN [6]
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RP MEDLINE=96206189; PubMed=8611530;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II) ADP.vanadate complex of the
RT Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";
RL Biochemistry 35:5404-5417(1996).
RN [7]
RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
RP MEDLINE=97452580; PubMed=9305951;
RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RT "X-ray structures of the MgADP, MgATPgammass, and MgAMPNP complexes
RT of the Dictyostelium discoideum myosin motor domain.";
RL Biochemistry 36:11619-11628(1997).
RN [8]
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RP MEDLINE=98070605; PubMed=9405148;
RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
RT "X-ray crystal structure and solution fluorescence characterization
RT of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the
RT Dictyostelium discoideum myosin motor domain.";
RL J. Mol. Biol. 274:394-407(1997).
RN CC
RN CC -!- FUNCTION: Myosin is a protein that binds to actin and has ATPase
RN CC activity that is activated by actin.
RN CC -!- SUBUNIT: Myosin II heavy chain is two-headed. It self-assembles
RN CC into filaments. Hexamer of 2 heavy chain subunits (MHC), 2 alkali
RN CC light chain subunits (MLC) and 2 regulatory light chain subunits
RN CC (MLC-2).
RN CC -!- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL

```



```

RC STRAIN=B204;
RX MEDLINE=21101013; PubMed=11159958;
RA Hsu T., Hutto D.L., Minion F.C., Wuerner R.L., Wannemuehler M.J.;
RT "Cloning of a beta-hemolysin gene of Brachyspira (Serpulina)
RT hyodysenteriae and its expression in Escherichia coli.";
RL Infect. Immun. 69:706-711(2001).
CC -!- FUNCTION: Carrier of the growing fatty acid chain in fatty acid
CC biosynthesis (By similarity). Has hemolytic activity forming pores
CC approximately 1 nm in diameter into erythrocytes. Is able to
CC induce murine colonic lesions and to disrupt the integrity of
CC epithelial cell monolayers.
CC -!- PATHWAY: De novo fatty acid biosynthesis.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: 4'-phosphopantetheine is transferred from CoA to a specific
CC serine of apo-ACP by acps. This modification is essential for
CC activity because fatty acids are bound in thioester linkage to the
CC sulphydryl of the prosthetic group (By similarity).
CC -!- SIMILARITY: Contains 1 acyl carrier domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U94886; AB868774.1; -.
CC DR HSSP; P02901; IACP.
CC DR HAMAP; MF_01217; -.
CC DR InterPro; IPR003231; Acyl_carrier.
CC DR InterPro; IPR006163; Pp_bind.
CC DR InterPro; IPR006162; Pantne_S.
CC DR Pfam; PF00550; pp-binding; 1.
CC DR ProDom; PD000887; Acyl_carrier; 1.
CC DR TIGRFAMs; TIGR00517; acyl_carrier; 1.
CC DR PROSITE; PS00075; ACP_DOMAIN; 1.
CC DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
CC KW Lipid synthesis; Fatty acid biosynthesis; Phosphopantetheine;
CC Virulence; Hemolysis; Toxin.
CC FT PROPEP 1 7
CC FT CHAIN 8 78
CC FT BINDING 37 37
CC FT BINDING 37 37
CC SQ SEQUENCE 78 AA; 8934 MW; 5F21E7ED0102C715 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 78;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 IEHKK 79
Db 73 IEHKK 78
|||||

RESULT 12
COAB.BPIKE STANDARD; PRT; 82 AA.
AC P03620;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coat protein B precursor (Major coat protein).
GN Villi.
OS Bacteriophage IKe.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
CC NCBI_TaxID=10867;
RN [1]_
CC SEQUENCE FROM N.A.
RX MEDLINE=85160831; PubMed=3981635;
RA Peeters B.P.H., Peters R.M., Schoenmakers J.G.G., Konings R.N.H.;
RT "Nucleotide sequence and genetic organization of the genome of the N-
RT specific filamentous bacteriophage IKe. Comparison with the genome of
RT the F-specific filamentous phages M13, fd and f1.";
```

```

RL J. Mol. Biol. 181:27-39(1985).
RN [2]
RP SEQUENCE OF 30-82.
RX MEDLINE=81215498; PubMed=7240173;
RA Nakashima Y., Frangione B., Wiseman R.L., Konigsberg W.H.;
RT "Primary structure of the major coat protein of the filamentous
RT bacterial viruses, If1 and IKe.";
RL J. Biol. Chem. 256:5792-5797(1981).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (5.0 ANGSTROMS).
RX MEDLINE=94118269; PubMed=8289247;
RA Marvin D.A., Hale R.D., Nave C., Citterich M.H.;
RT "Molecular models and structural comparisons of native and mutant
RT class I filamentous bacteriophages Ff (fd, f1, M13), If1 and IKe.";
RL J. Mol. Biol. 235:260-285(1994).
CC -!- FUNCTION: Coat protein B is the major coat protein of the virion.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X02139; CAA36072.1; -.
CC DR PIR; A92912; VCBPTK.
CC DR PDB; 1IEF; 3I-JUL-94.
CC DR InterPro; IPR008033; Phage_coat_Gp8.
CC DR Pfam; PF05371; Phage_Coat_Gp8; 1.
CC KW Coat protein; Signal; 3D-structure.
CC FT SIGNAL 1 29
CC FT CHAIN 30 82
CC FT CHAIN 31 82
CC FT HELIX 31 82
CC SQ SEQUENCE 82 AA; 8570 MW; C8F24F55A053799A CRC64;

Query Match 4.2%; Score 6; DB 1; Length 82;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 AKNKIA 114
Db 10 AKNKIA 15
|||||

RESULT 13
IMM1.PSEAE STANDARD; PRT; 87 AA.
AC Q06578;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Immunity protein for pyocin S1.
GN IMM1.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
CC Pseudomonadaceae; Pseudomonas.
CC NCBI_TaxID=287;
RN [1]_
CC SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RP STRAIN=NTH-H;
RX MEDLINE=9259934; PubMed=8491711;
RA Sano Y., Matsui H., Kobayashi M., Kageyama M.;
RT "Molecular structures and functions of pyocins S1 and S2 in
RT Pseudomonas aeruginosa.";
RL J. Bacteriol. 175:2907-2916(1993).
CC -!- SIMILARITY: BELONGS TO THE COLICINS COLE2/COLE8/COLE9 AND PYOCINS
CC S1/S2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: D12707; BAA02202.1; --  
 DR PIR: B36907; B36907.  
 DR HSSP: Q03708; 1CE1.  
 DR InterPro: IPR000290; Colicin\_pyocin.  
 DR Pfam: PF01320; Colicin\_Pyocin; 1.  
 DR PRINTS: PR01299; PYOCIN.  
 DR ProDom: PD007225; Colicin\_pyocin; 1.  
 DR KW Bacteriocin immunity.  
 SQ SEQUENCE 87 AA; 10040 MW; D332D0FD48556A6C CRC64;

Query Match 4.2%; Score 6; DB 1; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 TNNKKK 122  
 |||||  
 DB 22 TNNKKK 27

RESULT 14

IMM2\_PSEAE STANDARD; PRT; 87 AA.  
 AC Q06579;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Immunity protein for pyocin S2.  
 GN IMM2 OR PA1151.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=287;  
 [1]  
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.  
 RC STRAIN=PAO;  
 RX MEDLINE=93259934; PubMed=8491711;  
 RA Sano Y., Matsui H., Kobayashi M., Kageyama M.;  
 RT "Molecular structures and functions of pyocins S1 and S2 in  
 RT Pseudomonas aeruginosa."  
 RL J. Bacteriol. 175:2907-2916(1993).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PAO1;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Huynh W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brady L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an  
 RT opportunistic pathogen."  
 RL Nature 406:959-964(2000).  
 CC -1- SIMILARITY: BELONGS TO THE COLICINS COLE2/COLE8/COLE9 AND PYOCINS  
 CC S1/S2 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: D12708; BAA02204.1; --  
 DR EMBL: AE004545; AG04540.1; --  
 DR PIR: D36907; D36907.  
 DR HSSP: Q03708; 1CE1.  
 DR InterPro: IPR000290; Colicin\_pyocin.

DR Pfam: PF01320; Colicin\_Pyocin; 1.  
 DR PRINTS: PR01299; PYOCIN.  
 DR ProDom: PD007225; Colicin\_pyocin; 1.  
 DR KW Bacteriocin immunity; Complete proteome.  
 SQ SEQUENCE 87 AA; 10039 MW; DE9965FD485D424C CRC64;

Query Match 4.2%; Score 6; DB 1; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 TNNKKK 122  
 |||||  
 DB 22 TNNKKK 27

RESULT 15

Y031\_CHLTR STANDARD; PRT; 100 AA.  
 AC Q84034;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein CT031.  
 GN CT031.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=D/UW-3/Cx;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis."  
 RL Science 282:754-759(1998).  
 CC -1- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0121/CT031/TC0300  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: AE001277; AAC67621.1; --  
 DR PIR: B71567; B71567.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 100 AA; 11680 MW; 2A12B8A53EAA4F3C CRC64;

Query Match 4.2%; Score 6; DB 1; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 AKNKIA 114  
 |||||  
 DB 30 AKNKIA 35

RESULT 16

YGRB\_BACPF STANDARD; PRT; 108 AA.  
 ID YGRB\_BACPF  
 AC Q45130;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical 12.6 kDa protein in grpa 5' region (ORF6).  
 OS Bacillus pseudofirmus.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=79885;

```

[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=OF4;
RA MEDLINE=96198179; PubMed=8626304;
RA Sturr M.G., Krulwich T.A., Hicks D.B.;
RT "Purification of a cytochrome bd terminal oxidase encoded by the
RT Escherichia coli app locus from a delta cyo delta cyd strain
RT complemented by genes from Bacillus firmus OF4.";
RL J. Bacteriol. 178:1742-1749(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U39410; AAB05370.1; -.
DR Hypothetical protein; Transmembrane.
KW TRANSMEM 36 56 POTENTIAL.
FT TRANSMEM 58 78 POTENTIAL.
FT TRANSMEM 88 108 POTENTIAL.
SQ SEQUENCE 108 AA; 12555 MW; 12DC1461867CE513 CRC64;
Query Match 4.2%; Score 6; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 FKELEG 66
| | | | |
Db 5 FKELEG 10

RESULT 17
RS11 STRAW STANDARD; PRY; 134 AA.
ID ID RS11 STRAW STANDARD; PRY; 134 AA.
AC Q82DM3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 30S ribosomal protein S11.
GN RPSK OR SAV4952.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
OX [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RC MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Iehikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RA "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
CC [2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RC MEDLINE=22688305; PubMed=12652562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RA "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
CC -1- FUNCTION: Located on the platform of the 30S subunit, it bridges
CC several disparate RNA helices of the 16S rRNA. Forms part of the
CC Shine-Dalgarno cleft in the 70S ribosome (By similarity).
CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Interacts with
CC proteins S7 and S18. Binds to IF-3 (By similarity).
CC

```

```

CC -1- SIMILARITY: Belongs to the Sllp family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR HMAP; AF005041; BAC72664.1; -.
DR HMAP; MF 01310; -.
DR InterPro; IPR001971; Ribosomal_S11.
DR Pfam; PF00411; Ribosomal_S11; 1.
DR ProDom; PD001010; Ribosomal_S11; 1.
DR PROSITE; PS00054; RIBOSOMAL_S11; 1.
KW Ribosomal protein; RNA-binding; rRNA-binding; Complete proteome.
SQ SEQUENCE 134 AA; 14384 MW; 82DE7DB9D0D48982 CRC64;
Query Match 4.2%; Score 6; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 128 AKKVR 133
Db 10 AKKVR 15
RESULT 18
RS11 STRCO
ID RS11 STRCO STANDARD; PRT; 134 AA.
AC P72403;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 30S ribosomal protein S11.
GN RPSK OR SCO4728 OR SC634.06.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID:1902;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RC MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Ouail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147 (2002).
RN [2]
SEQUENCE OF 105-134 FROM N.A.
RP STRAIN=A3(2) / M145;
RP MEDLINE=97105902; PubMed=8948650;
RX Cho E.-J., Bae J.-B., Kang J.-G., Roe J.-H.;
RA "Molecular analysis of RNA polymerase alpha subunit gene from
RT Streptomyces coelicolor A3(2).";
RT Nucleic Acids Res. 24:4565-4571 (1996).
CC -1- FUNCTION: Located on the platform of the 30S subunit, it bridges
CC several disparate RNA helices of the 16S rRNA. Forms part of the
CC Shine-Dalgarno cleft in the 70S ribosome (By similarity).
CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Interacts with
CC proteins S7 and S18. Binds to IF-3 (By similarity).
CC -1- SIMILARITY: Belongs to the Sllp family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AL939121; CAA20384.1; --  
 DR EMBL; X92107; CAA63079.1; --  
 DR PIR; T35557; T35557.  
 DR HANAP; MF\_001310; --; 1.  
 DR InterPro; IPR001971; Ribosomal\_S11.  
 DR Pfam; PF00411; Ribosomal\_S11; 1.  
 DR ProDom; PD001010; Ribosomal\_S11; 1.  
 DR PROSITE; PS00054; RIBOSOMAL\_S11; 1.  
 KW Ribosomal protein; RNA-binding; Complete proteome.  
 SQ SEQUENCE 134 AA; 14398 MW; 868B6DBEA7D3F9F5 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 134;  
 Best Local Similarity 100.0%; Pred.No. 74;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 AKVRR 133

DB 10 AKVRR 15

# RESULT 19

ID Y056 METJA STANDARD; PRT; 136 AA.  
 AC Q60365;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein MJ0056.  
 GN MJ0056.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]

SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 EX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii.";  
 RL Science 273:1058-1073(1996).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; U67453; AAB98036.1; --  
 DR PIR; H64306; H64306.  
 DR TIGR; MJ0056; --

DR InterPro; IPR002834; DUF120.  
 DR Pfam; PF01982; DUF120; 1.  
 DR ProDom; PD015839; DUF120; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 136 AA; 15685 MW; 9F58C831BB21418C CRC64;

Query Match 4.2%; Score 6; DB 1; Length 136;  
 Best Local Similarity 100.0%; Pred.No. 75;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 102 LKLDRE 107

DB 46 LKLDRE 51

# RESULT 20

RL16 LEPIN STANDARD; PRT; 137 AA.  
 ID RL16 LEPIN STANDARD; PRT; 137 AA.  
 AC Q9XD29;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE 50S ribosomal protein L16.  
 GN RPL16 OR RPLP OR LA0746.  
 OS Leptospira interrogans.  
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
 OX NCBI\_TaxID=173;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=Lai / Serogroup Icterohaemorrhagiae / Serovar lai;  
 RX MEDLINE=20088835; PubMed=10620683;  
 RA Zuerner R.L., Hartskeerl R.A., van de Kemp H., Bal A.E.;  
 RT "Characterization of the Leptospira interrogans S10-spc-alpha  
 operon.";  
 RT FEMS Microbiol. Lett. 182:303-308(2000).  
 RL [2]

SEQUENCE FROM N.A.  
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;  
 RX MEDLINE=22598143; PubMed=12712204;  
 RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,  
 RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,  
 RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,  
 RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,  
 RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,  
 RA Sait Gironi I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,  
 RA Xu J.-G., Zhao G.-P.;  
 RT "Unique physiological and pathogenic features of Leptospira  
 RT interrogans revealed by whole-genome sequencing.";  
 RL Nature 422:888-893(2003).

CC FUNCTION: This protein binds directly to 23S ribosomal RNA and is  
 CC located at the a site of the peptidyltransferase center (By  
 CC similarity).

CC -!- SIMILARITY: Belongs to the L16P family of ribosomal proteins.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AF115283; AAN40590.1; --  
 DR EMBL; AE011262; AAN47945.1; --  
 DR InterPro; IPR001114; Ribosomal\_L16.  
 DR Pfam; PF00252; Ribosomal\_L16; 1.  
 DR PRINTS; PR00060; RIBOSOMALL16.  
 DR TIGRfams; TIGR01164; rplP\_bact; 1.  
 DR PROSITE; PS00566; RIBOSOMAL\_L16\_1; 1.  
 DR PROSITE; PS00701; RIBOSOMAL\_L16\_2; 1.  
 KW Ribosomal protein; rRNA-binding; Complete proteome.  
 SQ SEQUENCE 137 AA; 15505 MW; 24F371E7450FE257 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 137;  
 Best Local Similarity 100.0%; Pred.No. 75;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 EETAKK 130

|||||

Db 111 BETAKK 116

RESULT 21

RL11 HELP

ID RL11 HELP STANDARD; PRT; 141 AA.

AC P56037;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE SOS ribosomal protein L11.

GN PRLK OR HP1202 OR JHP1125.

OS Helicobacter pylori (Campylobacter pylori), and

OS Helicobacter pylori J99 (Campylobacter pylori J99).

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

OC Helicobacteriaceae; Helicobacter.

OX NCBI\_TaxID=216, 85963;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=26695 / ATCC 700392;

RX MEDLINE=97394467; PubMed=9252185;

RA Tomb J.-F., White O., Karlavage A.R., Clayton R.A., Sutton G.G.,

RA Fleisemann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,

RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,

RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,

RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,

RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,

RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,

RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,

RA Venter J.C.;

RT "The complete genome sequence of the gastric pathogen Helicobacter

RT pylori.";

RL Nature 388:539-547(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=J99;

RX MEDLINE=99120557; PubMed=9923682;

RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,

RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,

RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,

RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,

RA Trust T.J.;

RT "Genomic sequence comparison of two unrelated isolates of the human

RT gastric pathogen Helicobacter pylori.";

RL Nature 397:176-180(1999).

CC -!- FUNCTION: This protein binds directly to 23S ribosomal RNA (By

CC similarity).

CC -!- PTM: One or more lysine residues are methylated (By similarity).

CC -!- SIMILARITY: Belongs to the L11P family of ribosomal proteins.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; A8000626; AAD08248.1; -.

CC EMBL; A8001540; AAD06703.1; -.

CC PIR; B64670; B64670.

CC HSSP; P29395; 1MMS.

CC TIGR; HP1202; -.

CC HAMAP; MF\_00736; -; 1.

CC InterPro; IPR000911; Ribosomal\_L11.

CC InterPro; IPR006519; Ribosomal\_L11bac.

CC Pfam; PF00238; Ribosomal\_L11; 1.

CC Pfam; PF03946; Ribosomal\_L11\_N; 1.

CC ProDom; PD001367; Ribosomal\_L11; 1.

CC SMART; SM00649; RL11; 1.

CC TIGRfam; TIGR01632; RL11\_bact; 1.

CC ProSite; PS00359; Ribosomal\_L11; 1.

CC Ribosomal protein; rRNA-binding; Methylation; Complete proteome.

SQ SEQUENCE 141 AA; 15329 MW; D1BE047F62F5FAAB CRC64;

Query Match 4.2%; Score 6; DB 1; Length 141;

Best Local Similarity 100.0%; Pred.No.77; 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0;

Qy 110 KNKIAK 115

Db 94 KNKIAK 99

|||||

RESULT 22

GSPG ECOLI

ID GSPG ECOLI STANDARD; PRT; 145 AA.

AC P41442;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Putative general secretion pathway protein G precursor (Protein

DE transport protein hofG).

GN HOFG OR HOPG OR B3328.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI\_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=95204361; PubMed=7896718;

RA Stojiljkovic I., Schoenher R., Kusters J.G.;

RT "Identification of the hofG gene, a component of Escherichia coli

RT K-12 type II export system, and its conservation among different

RT pathogenic Escherichia coli and Shigella isolates.";

RL J. Bacteriol. 177:1892-1895(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blatter F.R., Blunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

CC -!- FUNCTION: Involved in a general secretion pathway (GSP) for the

CC export of proteins (By similarity).

CC -!- SIMILARITY: BELONGS TO THE PULG/OUTG/XPSG/EXSG/XCPT FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; U20786; AAA69031.1; -.

CC EMBL; U18397; AAA58125.1; -.

CC EMBL; AB000409; AAC76353.1; -.

CC PIR; B56150; B56150.

CC EcoGene; EG12886; hofG.

CC InterPro; IPR000983; Bac\_GSPG.

CC InterPro; IPR002416; Bac\_GSPH.

CC InterPro; IPR001120; Prok\_N methyl\_S.

CC PRINTS; PR00813; BCTERIALGSPG.

CC PRINTS; PR00885; BCTERIALGSPH.

CC PROSITE; PS00409; PROKAR\_NTER\_METHYL; 1.

CC Transport; Methylation; Complete proteome.

FT PROPEP 1 9

FT CHAIN 10 145

FT PUTATIVE GENERAL SECRETION PATHWAY

FT PROTEIN G.

FT MOD RES 10 10 METHYLATION (By similarity).

FT PROSITE; PS00359; Ribosomal\_L11; 1.

FT Ribosomal protein; rRNA-binding; Methylation; Complete proteome.

SQ SEQUENCE 145 AA; 15905 MW; C8CD67D743810072 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 145;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPP 8  
 DTD LISMO STANDARD; PRT; 148 AA.  
 ID MSCL\_PSESM  
 AC Q87WB2;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Large-conductance mechanosensitive channel.  
 GN MSCL OR PSPT04641.  
 OS Pseudomonas syringae (pv. tomato).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=323;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DC3000;  
 RX MEDLINE=22834015; PubMed=12928499;  
 RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,  
 RA Gwinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,  
 RA Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,  
 RA Nelson W.C., Davidson T., Zafar N., Zhou L., Liu J., Yuan Q.,  
 RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,  
 RA Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,  
 RA Alfano J.R., Carinhour S., Chatterjee A.K., Delaney T.P.,  
 RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,  
 RA White O., Fraser C.M., Collier A.;  
 RT "The complete genome sequence of the Arabidopsis and tomato pathogen  
 Pseudomonas syringae pv. tomato DC3000.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).  
 CC -!- FUNCTION: Channel that opens in response to stretch forces in the  
 membrane lipid bilayer. May participate in the regulation of  
 osmotic pressure changes within the cell (By similarity).  
 CC -!- SUBUNIT: Homopentamer (By similarity).  
 CC -!- SIMILARITY: Belongs to the mscl family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AB016872; A058087.1; -;  
 DR TIGR; PSPT04641; -;  
 DR HAMAP; MF\_00115; -; 1.  
 DR InterPro; IPR001185; MS channel.  
 DR Pfam; PF01741; Mscl; 1.  
 DR PRINTS; PR01264; MECHCHANNEL.  
 DR PROSITE; PS01327; MSCL; 1.  
 KW Ionic channel; transmembrane; Inner membrane; Complete proteome.  
 FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 16 42 BY SIMILARITY.  
 FT DOMAIN 43 73 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 74 95 BY SIMILARITY.  
 FT DOMAIN 96 148 CYTOPLASMIC (POTENTIAL).  
 SQ SEQUENCE 148 AA; 15617 MW; 3DE4BDA862605A2A CRC64;

Query Match 4.2%; Score 6; DB 1; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPP 8  
 DTD LISMO STANDARD; PRT; 150 AA.  
 ID MSCL\_PSESM  
 AC Q93RD9;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE D-tyrosyl-tRNA(Tyr) deacylase (EC 3.1.1.-).  
 GN DTD OR LMO1522.  
 OS Listeria monocytogenes.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OC NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD / Serovar 1/2a;  
 RA Okada Y.;  
 RT "Listeria monocytogenes rel and flanking regions.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD-e / Serovar 1/2a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 RA Charbit B., Chetoui F., Couve E., de Daruvar A., Dehoux P.,  
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,  
 RA Madsen E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
 RT "Comparative genomics of Listeria species.";  
 RL Science 294:849-852(2001).  
 CC -!- FUNCTION: Hydrolyzes D-tyrosyl-tRNA(Tyr) into D-tyrosine and free  
 tRNA(Tyr). Could be a defense mechanism against a harmful effect  
 of D-tyrosine (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -!- SIMILARITY: Belongs to the DTD family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AB051847; BAB50671.1; -;  
 DR EMBL; AL581979; CAC99600.1; -;  
 DR PIR; AB1265; AB1265.  
 DR HSSP; P32147; 1JKE.  
 DR ListList; LMO01522; -;  
 DR HAMAP; MF\_00518; -; 1.  
 DR InterPro; IPR003732; DTYRNA\_deacyl.  
 DR Pfam; PF02580; Tyr\_deacylase; 1.  
 DR ProDom; PD005653; DTYRNA\_deacyl; 1.  
 DR TIGSFams; TIGR00256; TIGR00256; 1.  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 150 AA; 16533 MW; 50FDD38497B7017D CRC64;

Query Match 4.2%; Score 6; DB 1; Length 150;

Best Local Similarity 100.0%; Pred. No. 82;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AEAGFI 44  
 DTD LISMO STANDARD; PRT; 150 AA.  
 ID MSCL\_PSESM  
 AC Q93RD9;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE D-tyrosyl-tRNA(Tyr) deacylase (EC 3.1.1.-).  
 GN DTD OR LMO1522.  
 OS Listeria monocytogenes.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OC NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD / Serovar 1/2a;  
 RA Okada Y.;  
 RT "Listeria monocytogenes rel and flanking regions.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD-e / Serovar 1/2a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 RA Charbit B., Chetoui F., Couve E., de Daruvar A., Dehoux P.,  
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,  
 RA Madsen E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
 RT "Comparative genomics of Listeria species.";  
 RL Science 294:849-852(2001).  
 CC -!- FUNCTION: Hydrolyzes D-tyrosyl-tRNA(Tyr) into D-tyrosine and free  
 tRNA(Tyr). Could be a defense mechanism against a harmful effect  
 of D-tyrosine (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -!- SIMILARITY: Belongs to the DTD family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AB051847; BAB50671.1; -;  
 DR EMBL; AL581979; CAC99600.1; -;  
 DR PIR; AB1265; AB1265.  
 DR HSSP; P32147; 1JKE.  
 DR ListList; LMO01522; -;  
 DR HAMAP; MF\_00518; -; 1.  
 DR InterPro; IPR003732; DTYRNA\_deacyl.  
 DR Pfam; PF02580; Tyr\_deacylase; 1.  
 DR ProDom; PD005653; DTYRNA\_deacyl; 1.  
 DR TIGSFams; TIGR00256; TIGR00256; 1.  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 150 AA; 16533 MW; 50FDD38497B7017D CRC64;



```

Db      114 ABAGFI 119
RESULT 25
YCCA_BACP3 STANDARD; PRT; 150 AA.
ID      YCCA_BACP3
AC      P55815;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Hypothetical protein in ccca 5' region (Fragment).
OS      Bacillus PS3 (Thermophilic bacterium PS-3).
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX      NCBI_TaxID=2334;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93379042; PubMed=7916623;
RA      Fujiwara Y., Oka M., Hamamoto T., Sone N.;
RT      "Chromosome c-551 of the thermophilic bacterium PS3, DNA sequence and
RT      analysis of the mature cytochrome.";
RL      Biochim. Biophys. Acta 1144:213-219(1993).
CC      -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC      -!- SIMILARITY: TO B.SUBTILIS YPJC, YQFU AND YIYT.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC      The European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X63125; -; NOT ANNOTATED_CDS.
DR      InterPro; IPR003740; DUF161.
DR      Pfam; PF02588; DUF161; 1.
KW      Hypothetical protein; Transmembrane.
FT      NON_TER
FT      TRANSMEM 19 39 POTENTIAL.
SQ      SEQUENCE 150 AA; 16394 MW; AEF892DB2CF7447 CRC64;
Query Match 4.2%; Score 6; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      130 KVRRAI 135
Db      80 KVRRAI 85
RESULT 26
RS7_THETH STANDARD; PRT; 155 AA.
AC      P17591;
DT      01-AUG-1990 (Rel. 15, Created)
DT      01-FEB-1995 (Rel. 31, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      30S ribosomal protein S7.
GN      rPSG OR RPS7.
OS      Thermus thermophilus.
OC      Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC      Thermus.
OX      NCBI_TaxID=274;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=HB8 / ATCC 27634;
RX      MEDLINE=90301504; PubMed=2362824;
RA      Yakhnin A.V., Vorozheykina D.P., Matvienko N.I.;
RT      "Nucleotide sequence of the Thermus thermophilus HB8 rps12 and rps7
RT      genes coding for the ribosomal proteins S12 and S7.";
RL      Nucleic Acids Res. 18:3659-3659(1990).
RN      [2]
RP      SEQUENCE OF 1-49.
RX      STRAIN=HB8 / ATCC 27634;
MEDLINE=95045586; PubMed=7957245;
RA      Tsiboli P., Hefurth E., Choli T.;
RT      "Purification and characterization of the 30S ribosomal proteins from
RT      the bacterium Thermus thermophilus.";
RL      Eur. J. Biochem. 226:169-177(1994).
RN      [3]
RP      X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX      MEDLINE=97473002; PubMed=9331418;
RA      Wimberly B.T., White S.W., Ramakrishnan V.;
RT      "The structure of ribosomal protein S7 at 1.9-A resolution reveals a
RT      beta-hairpin motif that binds double-stranded nucleic acids.";
RL      Structure 5:1187-1198(1997).
RN      [4]
RP      X-RAY CRYSTALLOGRAPHY (5.5 ANGSTROMS) OF THE 30S SUBUNIT.
RX      MEDLINE=99404610; PubMed=10476960;
RA      Clemens W.M. Jr., May J.L.C., Wimberly B.T., McCutcheon J.P.;
RT      "Structure of a bacterial 30S ribosomal subunit at 5.5 A resolution.";
RL      Nature 400:833-840(1999).
RN      [5]
RP      X-RAY CRYSTALLOGRAPHY (4.5 ANGSTROMS).
RX      MEDLINE=20056231; PubMed=10588692;
RA      Tocilj A., Schlutzen F., Janell D., Gluehmann M., Hansen H.A.,
RA      Harms J., Bashan A., Bartels H., Agmon I., Franceschi F., Yonath A.;
RT      "The small ribosomal subunit from Thermus thermophilus at 4.5 A
RT      resolution: pattern fittings and the identification of a functional
RT      site.";
RL      Proc. Natl. Acad. Sci. U.S.A. 96:14252-14257(1999).
RN      [6]
RP      X-RAY CRYSTALLOGRAPHY (3.05 ANGSTROMS) OF THE 30S SUBUNIT.
RX      MEDLINE=20466110; PubMed=11014182;
RA      Wimberly B.T., Brodersen D.E., Clemens W.M. Jr., Morgan-Warren R.J.,
RA      Carter A.P., Vornheim C., Hartsch T., Ramakrishnan V.;
RT      "Structure of the 30S ribosomal subunit.";
RL      Nature 407:327-339(2000).
RN      [7]
RP      X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS) OF THE 30S SUBUNIT.
RX      MEDLINE=20460546; PubMed=11007480;
RA      Schlutzen F., Tocilj A., Zariwach R., Harms J., Gluehmann M.,
RA      Janell D., Bashan A., Bartels H., Agmon I., Franceschi F., Yonath A.;
RT      "Structure of functionally activated small ribosomal subunit at 3.3
RT      A resolution.";
RL      Cell 102:615-623(2000).
RN      [8]
RP      X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS) OF THE 30S SUBUNIT.
RX      MEDLINE=21097713; PubMed=1163189;
RA      Brodersen D.E., Clemens W.M. Jr., Carter A.P., Morgan-Warren R.J.,
RA      Wimberly B.T., Ramakrishnan V.;
RT      "The structural basis for the action of the antibiotics tetracycline,
RT      pactamycin, and hygromycin B on the 30S ribosomal subunit.";
RL      Cell 103:1143-1154(2000).
RN      [9]
RP      X-RAY CRYSTALLOGRAPHY (3.00 ANGSTROMS) OF THE 30S SUBUNIT.
RX      MEDLINE=20466111; PubMed=11014183;
RA      Carter A.P., Clemens W.M. Jr., Brodersen D.E., Morgan-Warren R.J.,
RA      Wimberly B.T., Ramakrishnan V.;
RT      "Functional insights from the structure of the 30S ribosomal subunit
RT      and its interactions with antibiotics.";
RL      Nature 407:340-348(2000).
RN      [10]
RP      X-RAY CRYSTALLOGRAPHY (5.00 ANGSTROMS) OF THE RIBOSOME.
RX      MEDLINE=21402420; PubMed=11511350;
RA      Yusupova G.Z., Yusupov M.M., Cate J.H.D., Noller H.F.;
RT      "The path of messenger RNA through the ribosome.";
RL      Cell 106:233-241(2001).
RN      [11]
RP      X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS) OF THE 30S SUBUNIT.
RX      STRAIN=HB27;
RX      MEDLINE=21192023; PubMed=11296217;
RA      Pioletti M., Schlutzen F., Harms J., Zariwach R., Gluehmann M.,
RA      Avila H., Bashan A., Bartels H., Auerbach T., Jacobi C., Hartsch T.,
RA      Yonath A., Franceschi F.;
RT      "Crystals structures of complexes of the small ribosomal subunit with

```

RT tetracycline, edeine and IF3.;

RL ENBO J. 20:1829-1839(2001).

RN [12]

RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS) OF THE 30S SUBUNIT.

RX MEDLINE=21119514; PubMed=11228145;

RA Carter A.P., Clemons W.M. Jr., Brodersen D.E., Morgan-Warren R.J.,

RA Hartsch T., Wimberly B.T., Ramakrishnan V.;

RT "Crystal structure of an initiation factor bound to the 30S ribosomal

RT subunit.";

RT Science 291:498-501(2001).

RN [13]

RP X-RAY CRYSTALLOGRAPHY (5.5 ANGSTROMS) OF THE RIBOSOME.

RX MEDLINE=21238793; PubMed=11283358;

RA Yusupov M.M., Yusupova G.Z., Baucum A., Lieberman K., Earnest T.N.,

RA Cate J.H.D., Noller H.F.;

RT "Crystal structure of the ribosome at 5.5 A resolution.";

RL Science 292:883-896(2001).

RN [14]

RP X-RAY CRYSTALLOGRAPHY (3.11 ANGSTROMS) OF THE 30S SUBUNIT.

RX MEDLINE=21238794; PubMed=11340196;

RA Ogle J.M., Brodersen D.E., Clemons W.M. Jr., Tarry M.J., Carter A.P.,

RA Ramakrishnan V.;

RT "Recognition of cognate transfer RNA by the 30S ribosomal subunit.";

RL Science 292:897-902(2001).

RN [15]

RP X-RAY CRYSTALLOGRAPHY (3.05 ANGSTROMS) OF THE 30S SUBUNIT.

RX MEDLINE=21856772; PubMed=11866529;

RA Brodersen D.E., Clemons W.M. Jr., Carter A.P., Wimberly B.T.,

RA Ramakrishnan V.;

RT "Crystal structure of the 30S ribosomal subunit from Thermus

RT thermophilus: structure of the proteins and their interactions with

RT 16S RNA.";

RL J. Mol. Biol. 316:725-768(2002).

CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds

CC directly to 3' end of the 16S rRNA where it nucleates assembly of

CC the head domain of the 30S subunit. Is located at the subunit

CC interface close to the decoding center. Binds mRNA and the E-site

CC tRNA blocking its exit path in the ribosome.

CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S9

CC and S11. Binds to the C-terminus of IF3.

CC -!- SIMILARITY: Belongs to the S7P family of ribosomal proteins.

CC -!- DATABASE: NAME=Thermophilus 30S ribosomal subunit structure;

CC WWW="http://alfi.mrc-lmb.cam.ac.uk/-ribo/30S/".

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X52165; CAA36419.1; -.

CC PDB; 1RSS; 11-FEB-98.

CC PDB; 1DV4; 24-APR-00.

CC PDB; 1FKA; 04-SEP-00.

CC PDB; 1FJG; 25-SEP-00.

CC PDB; 1QD7; 17-JAN-01.

CC PDB; 1HR0; 24-JAN-01.

CC PDB; 1HNW; 21-FEB-01.

CC PDB; 1HNX; 21-FEB-01.

CC PDB; 1HNZ; 21-FEB-01.

CC PDB; 1I94; 12-APR-01.

CC PDB; 1I95; 12-APR-01.

CC PDB; 1I96; 12-APR-01.

CC PDB; 1I97; 12-APR-01.

CC PDB; 1GIX; 01-JUN-01.

CC PDB; 1IBK; 04-MAY-01.

CC PDB; 1IBL; 04-MAY-01.

CC PDB; 1IBM; 04-MAY-01.

CC PDB; 1JGO; 14-SEP-01.

CC PDB; 1JGP; 14-SEP-01.

CC PDB; 1JGQ; 14-SEP-01.

DR PDB; 1LIU; 22-MAR-02.

DR PDB; 1JSE; 12-APR-02.

DR PDB; 1N32; 29-NOV-02.

DR HAMAP; MF\_00480; -; 1.

DR InterPro; IPR000235; Ribosomal\_S7.

DR InterPro; IPR005717; Ribosomal\_S7\_b/o.

DR Pfam; PF00177; Ribosomal\_S7; 1.

DR ProDom; PD000817; Ribosomal\_S7; 1.

DR TIGRFAMs; TIGR01029; rpsG\_bact; 1.

DR PROSITE; PS00052; RIBOSOMAL\_S7; 1.

KW Ribosomal protein; rRNA-binding; rRNA-binding; rRNA-binding;

KW 3D-structure.

FT INIT\_MET 0

FT TURN 15 17

FT HELIX 20 29

FT TURN 32 33

FT HELIX 35 53

FT HELIX 57 68

FT STRAND 72 79

FT TURN 80 81

FT STRAND 82 89

FT HELIX 92 107

FT TURN 108 109

FT HELIX 115 127

FT TURN 128 129

FT HELIX 132 144

SQ SEQUENCE 155 AA; 17885 MW; BC405D68CBFE9474 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 155;

Best Local Similarity 100.0%; Pred. No. 84;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 ERMAEA 41

DB 141 ERMAEA 146

RESULT 27

RS7\_DEIRA STANDARD; PRT; 156 AA.

ID RS7 DEIRA

AC Q9RXK6;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 30S ribosomal protein S7.

GN RPSG OR DR0306.

OS Deinococcus radiodurans.

OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;

OC Deinococcaceae; Deinococcus.

OX NCBI\_TaxID=1299;

RV [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K1 / ATCC 13939 / DSM 20539 / NCIB 9279;

RX MEDLINE=20036896; PubMed=10567266;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,

RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,

RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,

RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,

RA Fraser C.M.;

RT "Genome sequence of the radioresistant bacterium Deinococcus

RT radiodurans R1.";

RL Science 286:1571-1577(1999).

CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds

CC directly to 16S rRNA where it nucleates assembly of the head

CC domain of the 30S subunit. Is located at the subunit interface

CC close to the decoding center, probably blocks exit of the E-site

CC tRNA (by similarity).

CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S9

CC and S11 (by similarity).

CC -!- SIMILARITY: Belongs to the S7P family of ribosomal proteins.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; AE001891; AAF09886.1; -  
 CC PIR; D75536; D75536.  
 CC HSP; P17291; LRSS.  
 CC TIGR; DR0305; -  
 CC HAMAP; MF\_00480; -; 1.  
 CC InterPro; IPR000235; Ribosomal\_S7.  
 CC InterPro; IPR005717; Ribosomal\_S7\_b/o.  
 CC Pfam; PF00177; Ribosomal\_S7; 1.  
 CC ProDom; PD000817; Ribosomal\_S7; 1.  
 CC TIGRFAMs; TIGR01029; rpsG\_bact; 1.  
 CC PROSITE; PS00052; RIBOSOMAL\_S7; 1.  
 CC Ribosomal protein; RNA-binding; rRNA-binding; tRNA-binding;  
 KW Complete proteome.  
 SQ SEQUENCE 156 AA; 17942 MW; C82E39470B6C4542 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 156;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 ERNVEA 41  
 DB 142 ERNVEA 147  
 |||||

RESULT 28  
 Y839\_HELHP  
 ID Y839\_HELHP STANDARD; PRT; 157 AA.  
 AC P59821;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Hypothetical UPF0090 protein H80839.  
 GN H80839.

OS Helicobacter hepaticus.  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC Helicobacteraceae; Helicobacter.  
 OX NCBI\_TaxID=32025;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 51449 / 3B1;  
 RX MEDLINE=22709201; PubMed=12810954;  
 RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,  
 RA Bell M., Droegge M., Partmann B., Fischer H.-P., Ge Z., Hoerster A.,  
 RA Holland R., Klein K., Koenig J., Macko L., Wendt G.L., Nyakatura G.,  
 RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;  
 RT "The complete genome sequence of the carcinogenic bacterium  
 RT Helicobacter hepaticus."  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).  
 CC -!- SIMILARITY: Belongs to the UPF0090 family.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; AS017146; AAP77436.1; -  
 CC HAMAP; MF\_01077; -; 1.  
 CC InterPro; IPR003728; DUF150.  
 CC Pfam; PF02576; DUF150; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 157 AA; 17843 MW; 574FB08D88473BCD CRC64;

Query Match 4.2%; Score 6; DB 1; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 IEQLAA 140  
 DB 10 IEQLAA 15  
 |||||

RESULT 29  
 UBC7\_WHEAT  
 ID UBC7\_WHEAT STANDARD; PRT; 168 AA.  
 AC P25868;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ubiquitin-conjugating enzyme E2 7 (EC 6.3.2.19) (Ubiquitin-protein  
 DE ligase 7) (Ubiquitin carrier protein 7).  
 GN UBC7.  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.; AND FUNCTION.  
 RC STRAIN=cv. Augusta;  
 RX MEDLINE=92052257; PubMed=1658801;  
 RA van Nocker S., Vierstra R.D.;  
 RT "Cloning and characterization of a 20-kDa ubiquitin carrier protein  
 RT from wheat that catalyzes multiubiquitin chain formation in vitro.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10297-10301(1991).  
 CC -!- FUNCTION: Catalyzes the covalent attachment of ubiquitin to other  
 CC proteins so as to signal them for selective protein degradation.  
 CC Involved in the formation of multiubiquitin chains.  
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
 CC diphosphate + protein N-ubiquityllysine.  
 CC -!- PATHWAY: Ubiquitin conjugation; second step.  
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
 CC thiolester formation (By similarity).  
 CC -!- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; M74077; -; NOT\_ANNOTATED\_CDS.  
 CC PIR; A41547; A41547.  
 CC HSP; Q02159; 2UCZ.  
 CC InterPro; IPR000608; UBQ\_conjugat.  
 CC Pfam; PF00179; UBQ\_con; 1.  
 CC ProDom; PD000461; UBQ\_conjugat; 1.  
 CC SMART; SM00212; UBQC; 1.  
 CC PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
 CC PROSITE; PS00127; UBIQUITIN\_CONJUGAT\_2; 1.  
 KW Ub1 conjugation pathway; Ligase; Multigene family.  
 FT BINDING 92 92 UBIQUITIN (BY SIMILARITY).  
 SQ SEQUENCE 168 AA; 18897 MW; 6ED5127DCAB415E2 CRC64;

RESULT 30

```

HSLV SHEON
ID HSLV SHEON STANDARD; PRT; 173 AA.
AC Q8E9V0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE ATP-dependent protease hslv (EC 3.4.25.-).
CN HSLV OR S04162
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
ON NCBI_TaxID=70863;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Imbraim M., Lee K., Barry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilsen K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
CC -!- FUNCTION: Protease subunit of a proteasome-like degradation
CC complex (By similarity).
CC -!- SUBUNIT: A double ring-shaped homohexamer of hslv is capped on
CC each side by a ring-shaped hslu homohexamer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family T1B. Hslv subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF015849; AAN57135.1; -.
CC TIGR; SO4162; -.
CC DR HAMAP; MF_00248; -.
CC DR InterPro; IPR001353; Peptidase T1.
CC DR Pfam; PF00227; Proteasome; 1.
CC KW Hydrolase; Protease; Threonine protease; Complete proteome.
CC INIT MET 0 0 BY SIMILARITY.
CC FT ACT_SITE 1 1 BY SIMILARITY.
CC SQ SEQUENCE 173 AA; 18786 MW; D31F5899C04783F2 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 173;
Best Local Similarity 100.0%; Pred.No. 93;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 AKKVR 133
DB 31 AKKVR 36

RESULT 31
HSLV YERPE
ID HSLV YERPE STANDARD; PRT; 173 AA.
AC Q8ZJ04;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE ATP-dependent protease hslv (EC 3.4.25.-).
CN HSLV OR YP00106 OR Y0295.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.

HSLV SHEON
ID HSLV SHEON STANDARD; PRT; 173 AA.
AC Q8E9V0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE ATP-dependent protease hslv (EC 3.4.25.-).
CN HSLV OR S04162
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
ON NCBI_TaxID=70863;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Imbraim M., Lee K., Barry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilsen K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
CC -!- FUNCTION: Protease subunit of a proteasome-like degradation
CC complex (By similarity).
CC -!- SUBUNIT: A double ring-shaped homohexamer of hslv is capped on
CC each side by a ring-shaped hslu homohexamer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family T1B. Hslv subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF015849; AAN57135.1; -.
CC TIGR; SO4162; -.
CC DR HAMAP; MF_00248; -.
CC DR InterPro; IPR001353; Peptidase T1.
CC DR Pfam; PF00227; Proteasome; 1.
CC KW Hydrolase; Protease; Threonine protease; Complete proteome.
CC INIT MET 0 0 BY SIMILARITY.
CC FT ACT_SITE 1 1 BY SIMILARITY.
CC SQ SEQUENCE 173 AA; 18786 MW; D31F5899C04783F2 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 173;
Best Local Similarity 100.0%; Pred.No. 93;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 AKKVR 133
DB 31 AKKVR 36

RESULT 32
HSLV PSEPK
ID HSLV PSEPK STANDARD; PRT; 175 AA.
AC Q88D28;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE ATP-dependent protease hslv (EC 3.4.25.-).
CN HSLV OR PP5000.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.

```

```

OX NCBI_TaxID=160488;
RN SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holzapple E., Scallan D., Tran K., Moazzes A.,
RA Uterback T., Rizzo M., Lee K., Kosack D., Moesti D., Wedler H.,
RA Lauber J., Stjepandic D., Hohesl J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuermler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
CC -!- FUNCTION: Protease subunit of a proteasome-like degradation
CC complex (By similarity).
CC -!- SUBUNIT: A double ring-shaped homohexamer of hslv is capped on
CC each side by a ring-shaped hslu homohexamer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family T1B. Hslv subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AE016792; AAN70566.1; -.
DR TIGR; PP5000; -.
DR HAMAP; MF_00248; -.
DR InterPro; IPR001353; Peptidase_T1.
DR Pfam; PF00227; Proteasome; 1.
KW Hydrolase; Protease; Threonine protease; Complete proteome.
FT INIT MET 0 0 BY SIMILARITY.
FT ACT SITE 1 1 BY SIMILARITY.
SQ SEQUENCE 175 AA; 18635 MW; 7789B55A54152B8B CRC64;

Query Match 4.2%; Score 6; DB 1; Length 175;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 AKKVR 133
Db |||||
31 AKKVR 36

RESULT 33
HSLV_PSESM STANDARD; PRT; 175 AA.
AC Q87V01;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ATP-dependent protease hslv (EC 3.4.25.-).
GN HSLV OR PSPT05140.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499;
RA Buell C.R., Jørdar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S., Brinkac L., Reanan M.J., Haft D.H.,
RA Nelson W.C., Davidson T., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,
RA Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,

```

```

RA Alfano J.R., Cartinhour S., Chatterjee A.K., Delaney T.P.,
RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
RA White O., Fraser C.M., Collmer A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
CC -!- FUNCTION: Protease subunit of a proteasome-like degradation
CC complex (By similarity).
CC -!- SUBUNIT: A double ring-shaped homohexamer of hslv is capped on
CC each side by a ring-shaped hslu homohexamer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family T1B. Hslv subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AE016874; AAC58567.1; -.
DR TIGR; PSPT05140; -.
DR HAMAP; MF_00248; -.
DR InterPro; IPR001353; Proteasome; 1.
DR Pfam; PF00227; Proteasome; 1.
KW Hydrolase; Protease; Threonine protease; Complete proteome.
FT INIT MET 0 0 BY SIMILARITY.
FT ACT SITE 1 1 BY SIMILARITY.
SQ SEQUENCE 175 AA; 18695 MW; 8DBF852C017E80C2 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 175;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 AKKVR 133
Db |||||
31 AKKVR 36

RESULT 34
IPYR_ECOL6 STANDARD; PRT; 175 AA.
AC Q8FAG0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
DE hydrolase) (Ppase).
GN PPA OR CS323.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6.H1 / CFT073 / ATCC 700928;
RX MEDLINE=22382334; PubMed=12471157;
RA Welch R.A., Surland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Ferna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.
CC -!- COFACTOR: Binds 4 magnesium ions per subunit. Other metal ions can
CC support activity, but at a lower rate. Two magnesium ions are
CC required for the activation of the enzyme and are present before
CC substrate binds, two additional magnesium ions form complexes with
CC substrate and product (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the Ppase family.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE016771; AAN83744.1; --  
 DR HAMAP: MF\_00209; 1;  
 DR InterPro: IPR008163; Inorg\_pphsph.  
 DR InterPro: IPR008162; Pyrophosphatase.  
 DR Pfam: PF00719; Pyrophosphatase; 1.  
 DR ProDom: PD002014; Inorg\_pphsph; 1.  
 DR PROSITE: PS00387; PPASE; 1.  
 KW Hydrolase; Metal-binding; Magnesium; Complete proteome.  
 FT INIT MET 0  
 FT METAL 65 65 BY SIMILARITY.  
 FT METAL 70 70 MAGNESIUM 1 (BY SIMILARITY).  
 FT METAL 102 102 MAGNESIUM 1 (BY SIMILARITY).  
 SQ SEQUENCE 175 AA; 19600 MW; 7DBAE4D08DAD8FDB CRC64;

Query Match 4.2%; Score 6; DB 1; Length 175;

Best Local Similarity 100.0%; Pred. No. 94; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0;

Qy 107 ERAKXK 112

Db 170 ERAKXK 175

RESULT 35

ID IPYR\_ECOLI STANDARD; PRT; 175 AA.  
 AC P17288;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-  
 DE hydrolase) (Pase).  
 DE PPA OR B4226 OR Z5837 OR RCS5204.  
 OS Escherichia coli; and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562, 83334;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=K12;  
 RX MEDLINE=89053923; PubMed=2848015;  
 RA Lahti R., Pitkaranta T., Valve E., Ilta I., Kukko-Kalske E.,  
 RA Heinonen J.;  
 RT "Cloning and characterization of the gene encoding inorganic  
 RT pyrophosphatase of Escherichia coli K-12.";  
 RL J. Bacteriol. 170:5901-5907(1988).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=K12 / MG1655;  
 RX MEDLINE=95334362; PubMed=7610040;  
 RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,  
 RA Blattner F.R.;  
 RT "Analysis of the Escherichia coli genome VI: DNA sequence of the  
 RT region from 92.8 through 100 minutes.";  
 RL Nucleic Acids Res. 23:2105-2119(1995).  
 RN [3]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamasis K.,

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN [4]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 RN [5]  
 RC SEQUENCE OF 1-20 AND 94-105.  
 RP STRAIN=K12 / EMG2; PubMed=9298646;  
 RX MEDLINE=97443975; PubMed=9298646;  
 RA Link A.J., Robison K., Church G.M.;  
 RT "Comparing the predicted and observed properties of proteins encoded  
 RT in the genome of Escherichia coli K-12.";  
 RL Electrophoresis 18:1259-1313(1997).  
 RN [6]  
 RC SIMILARITY TO YEAST AND K.LACTIS PPASES.  
 RP MEDLINE=90254161; PubMed=2160278;  
 RX Lahti R., Kolaowski L.P. Jr., Heinonen J., Vihinen M., Pohjanoksa K.,  
 RA Cooperman B.S.;  
 RT "Conservation of functional residues between yeast and E. coli  
 RT inorganic pyrophosphatases.";  
 RL Biochim. Biophys. Acta 1038:338-345(1990).  
 RN [7]  
 RC MUTAGENESIS.  
 RP MEDLINE=90344799; PubMed=1974462;  
 RX Lahti R., Pohjanoksa K., Pitkaranta T., Heikinheimo P., Salminen T.,  
 RA Meyer P., Heinonen J.;  
 RT "A site-directed mutagenesis study on Escherichia coli inorganic  
 RT pyrophosphatase. Glutamic acid-98 and lysine-104 are important for  
 RT structural integrity, whereas aspartic acids-97 and -102 are  
 RT essential for catalytic activity.";  
 RL Biochemistry 29:5761-5766(1990).  
 RN [8]  
 RC MUTAGENESIS OF TYROSINE RESIDUES.  
 RP MEDLINE=91249822; PubMed=1645654;  
 RX Lahti R., Salminen T., Latonen S., Heikinheimo P., Pohjanoksa K.,  
 RA Heinonen J.;  
 RT "Genetic engineering of Escherichia coli inorganic pyrophosphatase.  
 RT Tyr55 and Tyr141 are important for the structural integrity.";  
 RL Eur. J. Biochem. 198:293-297(1991).  
 RN [9]  
 RC MODIFICATION OF TYR-149.  
 RP MEDLINE=93185883; PubMed=8383066;  
 RX Kaneko S., Ichiba T., Hirano N., Hachimori A.;  
 RT "Modification of tryptophan 149 of inorganic pyrophosphatase from  
 RT Escherichia coli.";  
 RL Int. J. Biochem. 25:233-238(1993).  
 RN [10]  
 RC X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
 RP MEDLINE=95062129; PubMed=7971944;  
 RX Kankare J., Neal G.S., Salminen T., Glumhoff T., Cooperman B.S.,  
 RA Lahti R., Goldman A.;  
 RT "The structure of E.coli soluble inorganic pyrophosphatase at 2.7-A  
 RT resolution.";  
 RL Protein Eng. 7:823-830(1994).  
 RN [11]  
 RC ERRATUM.  
 RP Kankare J., Neal G.S., Salminen T., Glumhoff T., Cooperman B.S.,  
 RA Lahti R., Goldman A.;  
 RL Protein Eng. 7:1173-1173(1994).  
 RN [12]  
 RC X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RP MEDLINE=96267057; PubMed=8664256;  
 RX Kankare J., Salminen T., Lahti R., Cooperman B.S., Baykov A.A.,

RA Goldman A.;  
RT "Crystallographic identification of metal-binding sites in  
RN Escherichia coli inorganic pyrophosphatase.";  
RL Biochemistry 35:4670-4677(1996).  
[13]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Karkare J., Salminen T., Lahti R., Cooperman B.S., Baykov A.A.,  
RA Goldman A.;  
RT "Structure of Escherichia coli inorganic pyrophosphatase at 2.2-A  
RT resolution.";  
RL Acta Crystallogr. D 52:551-563(1996).  
[14]  
RN X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS).  
RP MEDLINE=97379370; PubMed=9237692;  
RX Aavaeva S., Kurilova S., Nazarova T., Rodina E., Vorobyeva N.,  
RA Sklyankina K., Grigorjeva O., Hartutyunyan E., Oganessyan V.,  
RA Wilson K., Dauter Z., Huber R., Mather T.;  
RT "Crystal structure of Escherichia coli inorganic pyrophosphatase  
RT complexed with SO4(2-). Ligand-induced molecular asymmetry.";  
RL FEBS Lett. 410:502-508(1997).  
[15]  
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RP MEDLINE=97352630; PubMed=9201917;  
RX Hartutyunyan E.H., Oganessyan V.Y., Oganessyan N.N., Aavaeva S.M.,  
RA Nazarova T.I., Vorobyeva N.N., Kurilova S.A., Huber R., Mather T.;  
RT "Crystal structure of holo inorganic pyrophosphatase from Escherichia  
RT coli at 1.9 A resolution. Mechanism of hydrolysis.";  
RL Biochemistry 36:7754-7760(1997).  
CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.  
CC -!- COFACTOR: Binds 4 magnesium ions per subunit. Other metal ions can  
CC support activity, but at a lower rate. Two magnesium ions are  
CC required for the activation of the enzyme and are present before  
CC substrate binds, two additional magnesium ions form complexes with  
CC substrate and product.  
CC -!- SUBUNIT: Homohexamer.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the PPase family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M23550; AAB88583.1; -;  
DR EMBL; U14003; AAA97123.1; -;  
DR EMBL; AB000494; AAC77183.1; -;  
DR EMBL; AB003655; AAG59424.1; -;  
DR EMBL; AF002568; BAB38627.1; -;  
DR PIR; A27648; PWEC.  
DR PIR; D86120.  
DR PIR; D91279; D91279.  
DR PDB; 1FAJ; 08-NOV-96.  
DR PDB; 1IGF; 20-DEC-94.  
DR PDB; 1INO; 03-APR-96.  
DR PDB; 1IPW; 20-AUG-97.  
DR PDB; 1JFD; 03-DEC-97.  
DR PDB; 1MWJ; 03-DEC-97.  
DR PDB; 1MXJ; 03-DEC-97.  
DR PDB; 1MYJ; 03-DEC-97.  
DR PDB; 1MUZ; 03-DEC-97.  
DR PDB; 1OBW; 04-SEP-97.  
DR PDB; 2EIP; 08-NOV-96.  
DR SWISS-2DPAGE; P17288; COLI.  
DR EcoGene; EG10755; ppa.  
DR HAMAP; MF 00209; -; 1.  
DR InterPro; IPR008163; Inorg\_pphsh.  
DR InterPro; IPR008162; Pyrophosphatase.  
DR Pfam; PF00719; Pyrophosphatase; 1.  
DR ProDom; PD002014; Inorg\_pphsh; 1.  
DR PROSITE; PS00387; PPASE; 1.

KW Hydrolase; Metal-binding; Magnesium; 3D-structure; Complete proteome.  
FT INIT\_MET 0  
FT METAL 65 65 MAGNESIUM 1.  
FT METAL 70 70 MAGNESIUM 1 AND 2.  
FT METAL 102 102 MAGNESIUM 1.  
FT MUTAGEN 20 20 E->D: 16% ACTIVITY.  
FT MUTAGEN 29 29 K->R: 2% ACTIVITY.  
FT MUTAGEN 31 31 E->D: 6% ACTIVITY.  
FT MUTAGEN 43 43 R->K: 10% ACTIVITY.  
FT MUTAGEN 51 51 Y->F: 64% ACTIVITY.  
FT MUTAGEN 55 55 Y->F: 7% ACTIVITY.  
FT MUTAGEN 65 65 D->E: 6% ACTIVITY.  
FT MUTAGEN 67 67 D->E: 1% ACTIVITY.  
FT MUTAGEN 70 70 D->E: NO ACTIVITY.  
FT MUTAGEN 97 97 D->E: 22% ACTIVITY.  
  
Query Match 4.2%; Score 6; DB 1; Length 175;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 107 ERAKNK 112  
| | | | |  
Db 170 ERAKNK 175  
  
RESULT 36  
HSLV PSEAE STANDARD; PRT; 176 AA.  
ID HSLV PSEAE  
AC QSHUC6; 2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE ATP-dependent protease hslv (EC 3.4.25.-).  
GN HSLV OR PA5053.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Brody L.L., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Smith K.A., Spencer D.H., Hancock R.E.W., Lory S., Olson M.V.;  
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an  
RT opportunistic pathogen.";  
RL Nature 406:959-964(2000).  
CC -!- FUNCTION: Protease subunit of a proteasome-like degradation  
CC complex (By similarity).  
CC -!- SUBUNIT: A double ring-shaped homohexamer of hslv is capped on  
CC each side by a ring-shaped hslu homohexamer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to peptidase family T1B. Hslv subfamily.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; AE004918; AAG08438.1; -;  
DR PIR; D83015; D83015.  
DR HSSP; P31059; 1B94.  
DR MEROPS; T01.006; -;  
DR HAMAP; MF\_00248; -; 1.  
DR InterPro; IPR001353; Peptidase\_T1.  
DR Pfam; PF00227; proteasome; 1.

KW Hydrolase; Protease; Threonine protease; Complete proteome.  
 FT INIT MET 0 BY SIMILARITY.  
 FT ACT SITE 1 BY SIMILARITY.  
 SQ SEQUENCE 176 AA; 18654 MW; 291076FF451B5B09 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 94; 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0;

QY 128 AKVRR 133  
 Db 31 AKVRR 36  
 |||||

RESULT 37  
 YGV5 YEAST STANDARD; PRT; 178 AA.  
 AC P53071;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Hypothetical 19.3 kDa protein in HAP2-ADE5,6 intergenic region.  
 GN YGL235W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Farman B., Kramer W.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; 272758; CAA96954.1; -;  
 DR PIR; S64257;  
 DR GerOnline; 141284; -;  
 DR SGD; S0003204; YGL235W.  
 KW Hypothetical protein.  
 SQ SEQUENCE 178 AA; 19334 MW; 2972A987E576A6E9 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 GCAPLS 88  
 Db 25 GCAPLS 30  
 |||||

RESULT 38  
 YJV7 YEAST STANDARD; PRT; 198 AA.  
 AC P40853;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Hypothetical 22.0 kDa protein in HX171-HX18 intergenic region.  
 GN YJL217W OR J0226 OR HRC198.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Farman B., Kramer W.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; 272758; CAA96954.1; -;  
 DR PIR; S64257;  
 DR GerOnline; 141284; -;  
 DR SGD; S0003204; YGL235W.  
 KW Hypothetical protein.  
 SQ SEQUENCE 178 AA; 19334 MW; 2972A987E576A6E9 CRC64;

RA Vandebol M., Durand P., Bolle P.-A., Dion C., Portetelle D.,  
 RA Hilger F.;  
 RT "Sequence analysis of a 40.2 kb DNA fragment located near the left  
 RT telomere of yeast chromosome X.";  
 RL Yeast 10:1657-1662(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; 234098; CAA83993.1; -;  
 DR EMBL; 249492; CAA89514.1; -;  
 DR PIR; S50710; S50710.  
 DR GerOnline; 141829; -;  
 DR SGD; S0003753; YJL217W.  
 KW Hypothetical protein.  
 SQ SEQUENCE 198 AA; 21967 MW; 52F6BB7CEA722D37 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 ACTPER 37  
 Db 169 ACTPER 174  
 |||||

RESULT 39  
 RNH2 FORGI STANDARD; PRT; 201 AA.  
 ID RNH2 FORGI  
 AC Q51832;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ribonuclease HII (EC 3.1.26.4) (RNase HII).  
 GN RNHB OR PG0736.  
 OS Porphyromonas gingivalis (Bacteroides gingivalis).  
 CC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 CC Porphyromonadaceae; Porphyromonas.  
 OX NCBI\_TaxID=837;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=W83;  
 RX MEDLINE=22829867; PubMed=12949112;  
 RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,  
 RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.,  
 RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T., Tallon L., Gray J.,  
 RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,  
 RA Dewhirst F.E., Fraser C.M.;  
 RT "Complete genome sequence of the oral pathogenic bacterium  
 RT Porphyromonas gingivalis strain W83.";  
 RL J. Bacteriol. 185:5591-5601(2003).

RN [2]  
 RP SEQUENCE OF 34-201 FROM N.A.

RC STRAIN=W83;  
 RA Rigg G.P., Roberts I.S.;  
 RT "Molecular analysis of PgaA an antigen from periodontopathogen  
 RT Porphyromonas gingivalis";  
 RT Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
 CC RNA-DNA hybrids specifically (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphomonoester.  
 CC -!- COFACTOR: Manganese (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -!- SIMILARITY: Belongs to the RNase HII family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -



CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE017174; RAQ65905.1; --

DR EMBL; X95938; CAAG5177.1; --

DR TIGR; PG0736; --

DR HAMAP; MF\_00052; --; 1.

DR InterPro; IPR001352; RNase HII/HIII.

DR Pfam; PF01351; RNase HII; 1.

KW Hydrolase; Nuclease; Endonuclease; Manganese; Complete proteome.

FT ACT\_SITE 17 17 BY SIMILARITY.

FT ACT\_SITE 109 109 BY SIMILARITY.

FT ACT\_SITE 128 128 BY SIMILARITY.

SQ SEQUENCE 201 AA; 22733 MW; 812CF1455A59B20C CRC64;

Query Match 4.2%; Score 6; DB 1; Length 201;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 RATEQL 138

DB 94 RATEQL 99

RESULT 40

ID TRI\_PONLE STANDARD; PRT; 201 AA.

AC P05547;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Troponin I.

OS leptoactylus (Narrow-fingered crayfish) (Astacus

leptoactylus).

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;

OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;

OC Astacoidea; Astacidae; Pontastacus.

OX NCBI\_TaxID=6717;

RN [1]

RP MEDLINE=89109165; PubMed=2912973;

RA Kobayashi T., Takagi T., Konishi K., Cox J.A.;

RT "Amino acid sequence of crayfish troponin I.";

RL J. Biol. Chem. 264:1551-1557(1989).

CC -!- FUNCTION: Troponin I is the actomyosin ATPase inhibitory subunit

CC present in the thin filament regulatory complex.

CC -!- MISCELLANEOUS: There is a 30 residue long N-terminal tail that

CC does not occur in skeletal muscle Tni's, but is present in cardiac

CC muscle Tni's.

CC -!- SIMILARITY: Belongs to the troponin I family.

DR InterPro; IPR001978; Troponin.

DR Pfam; PF00992; Troponin; 1.

KW Methylation; Actin-binding; Acetylation.

FT MOD\_RES 1 1 ACETYLATION (PROBABLE).

FT MOD\_RES 142 142 METHYLATION (TRI-).

FT MOD\_RES 146 146 METHYLATION (TRI-).

FT DOMAIN 108 117 TROPONIN T-INTERACTION.

FT DOMAIN 135 148 ACTIN-BINDING.

SQ SEQUENCE 201 AA; 23490 MW; 47585B56DB8A65 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 201;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 KKKEFE 125

DB 175 KKKEFE 180

RESULT 41

ZFP\_IRV6 STANDARD; PRT; 208 AA.

ID ZFP\_IRV6

AC P47732; O55770;

DT 01-FEB-1996 (Rel. 33, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Zinc finger protein.

OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).

CC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.

OX NCBI\_TaxID=10488;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92196996; PubMed=1549908;

RA Handermann M., Schnitzler P., Roesen-Wolff A.P., Raab K.,

RA Sontag K.C., Darai G.

RT "Identification and mapping of origins of DNA replication within the

RT DNA sequences of the genome of insect iridescent virus type 6.";

RL Virus Genes 6:19-32(1992).

RN [2]

RP REVISIONS TO C-TERMINUS.

RA Jacob N.J., Mueller K., Bahr U., Darai G.;

RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: Contains 1 BIR repeat.

CC -!- SIMILARITY: Contains 1 RING-type zinc finger.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; AF303741; AAB94481.1; --

DR HSSP; O15392; 1B31.

DR InterPro; IPR001370; BIR.

DR InterPro; IPR001841; Znf\_ring.

DR Pfam; PF00653; BIR; 1.

DR Pfam; PF00097; zf-C3HC4; 1.

DR SMART; SMO0238; BIR; 1.

DR SMART; SMO0184; RING; 1.

DR PROSITE; PS01282; BIR\_REPEAT\_1; FALSE\_NEG.

DR PROSITE; PS00143; BIR\_REPEAT\_2; 1.

DR PROSITE; PS00518; ZF\_RING\_1; FALSE\_NEG.

DR PROSITE; PS00089; ZF\_RING\_2; 1.

KW Zinc-finger; Repeat.

FT REPEAT 37 108 BIR.

FT DOMAIN 134 151 3 X 6 AA TANDEM REPEATS.

FT REPEAT 134 139 1.

FT REPEAT 140 145 2.

FT REPEAT 146 151 3.

FT ZN\_FING 163 197 RING-TYPE.

SQ SEQUENCE 208 AA; 24142 MW; 41A2E2FC18833390 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 208;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 BEHKKH 80

DB 92 BEHKKH 97

RESULT 42

CYB6\_EUGGR

ID CYB6\_EUGGR STANDARD; PRT; 214 AA.

AC P31480;

DT 01-JUL-1993 (Rel. 26, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Cytochrome b6.

GN PETE.

OS Euglena gracilis.

OG Chloroplast.  
 OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.  
 OX NCBI\_TaxID=3039;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Z / UTEX 753;  
 RX MEDLINE=95007823; PubMed=7923415;  
 RA Hong L., Hallick R.B.;  
 RT "Gene structure and expression of a novel Euglena gracilis  
 RT chloroplast operon encoding cytochrome b6 and the beta and epsilon  
 RT subunits of the H(+)-ATP synthase complex.";  
 RL Curr. Genet. 25:270-281(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Z / UTEX 753;  
 RX MEDLINE=93347989; PubMed=8346031;  
 RA Hallick R.B., Hong L., Drager R.G., Favreau M.R., Monfort A.,  
 RA Orsat B., Spielmann A., Stutz E.;  
 RT "Complete sequence of Euglena gracilis chloroplast DNA";  
 RL Nucleic Acids Res. 21:3537-3544(1993).  
 RN [3]  
 RP SEQUENCE OF 1-30, AND SUBCELLULAR LOCATION.  
 RX MEDLINE=22721363; PubMed=12837550;  
 RA Santillan Torres J.L., Atreia A., Claros M.G., Gonzalez-Halphen D.;  
 RT "Cytochrome f and subunit IV, two essential components of the  
 RT photosynthetic bc1 complex typically encoded in the chloroplast  
 RT genome, are nucleus-encoded in Euglena gracilis";  
 RL Biochim. Biophys. Acta 1604:180-189(2003).  
 CC -!- FUNCTION: Component of the cytochrome b6/f complex which is part  
 CC of the photosynthetic respiratory chain.  
 CC -!- COPACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or  
 CC b562) is low-potential and absorbs at about 562 nm, and heme 2 (or  
 CC BH or b566) is high-potential and absorbs at about 566 nm. 17  
 CC -!- SUBUNIT: The main subunits of complex b6-f are: cytochrome b6, 17  
 CC kba polypeptide (petb), cytochrome f and the Rieske protein.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast  
 CC thylakoid membrane (By similarity).  
 CC -!- SIMILARITY: Belongs to the cytochrome b family. PetB subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X70810; CAA50129.1; -  
 CC PIR; S34548; S34548.  
 CC HAMAP; MF\_00633; -; 1.  
 CC InterPro; IPR005797; Cytb\_b6\_N.  
 CC Pfam; PF00033; Cytochrome b\_N. 1.  
 CC PROSITE; PS00192; CYTOCHROME B HEME; 1.  
 CC TRANSPORT: Electron transport; Photosynthesis; Chloroplast; Thylakoid;  
 CC Heme; Metal-binding; Iron; Transmembrane.  
 FT INIT\_MET 0  
 FT TRANSFORM 31 51 POTENTIAL.  
 FT TRANSFORM 89 109 POTENTIAL.  
 FT TRANSFORM 115 135 POTENTIAL.  
 FT TRANSFORM 185 205 POTENTIAL.  
 FT METAL 85 85 IRON (HEME 1 AXIAL LIGAND) (BY  
 FT METAL 99 99 IRON (HEME 2 AXIAL LIGAND) (BY  
 FT METAL 186 186 IRON (HEME 1 AXIAL LIGAND) (BY  
 FT METAL 201 201 IRON (HEME 2 AXIAL LIGAND) (BY  
 FT SEQUENCE 214 AA; 24254 MW; B4CBAP57DDA2F4D CRC64;  
 Query Match 4.2%; Score 6; DB 1; Length 214;  
 Best Local Similarity 100.0%; Pred.No.1.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 85 AFLSVK 90  
 DB 64 AFLSVK 69  
 RESULT 43  
 ID DEF MYCGE STANDARD; PRT; 216 AA.  
 AC P47352;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).  
 OS Mycoplasma genitalium.  
 GN Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RT "The minimal gene complement of Mycoplasma genitalium";  
 RL Science 270:397-403(1995).  
 CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of  
 CC newly synthesized proteins. Requires at least a dipeptide for an  
 CC efficient rate of reaction. N-terminal L-methionine is a  
 CC prerequisite for activity but the enzyme has broad specificity at  
 CC other positions (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +  
 CC methionyl peptide.  
 CC -!- COPACTOR: Binds 1 iron(II) ion (By similarity).  
 CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U39691; AAC71324.1; ALT\_INIT.  
 CC HSP; P27251; 2DEF.  
 CC TIGR; MG106; -  
 CC HAMAP; MF\_00163; -; 1.  
 CC InterPro; IPR000181; Pep\_deformylase.  
 CC Pfam; PF01327; Pep\_deformylase; 1.  
 CC PRINTS; PR01576; PDEFORMYLASE.  
 CC ProDom; PD003844; Pep\_deformylase; 1.  
 CC TIGRFAMs; TIGR00079; pept\_deformyl; 1.  
 CC ACT\_SITE 179 179 BY SIMILARITY.  
 CC METAL 134 134 IRON (BY SIMILARITY).  
 CC METAL 178 178 IRON (BY SIMILARITY).  
 CC METAL 182 182 IRON (BY SIMILARITY).  
 CC SEQUENCE 216 AA; 25171 MW; 52D2E8FA58BEFC5 CRC64;  
 Query Match 4.2%; Score 6; DB 1; Length 216;  
 Best Local Similarity 100.0%; Pred.No.1.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 87 LSVKKQ 92  
 DB 135 LSVKKQ 140

```

RESULT 44
YG30 YEAST
ID YG30 YEAST STANDARD; PRT; 217 AA.
AC P48238;
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 25.2 kDa protein in RSRI-CYS4 intergenic region.
GN YGR153W OR G6661.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=96158062; PubMed=8585325;
RA Skala J., Nawrocki A., Goffeau A.;
RT "The sequence of a 27 kb segment on the right arm of chromosome VII
RT from Saccharomyces cerevisiae reveals MOLI, NAT2, RPL30B, RSRI, CYS4,
RT PEM1/CHO2, NSRI genes and ten new open reading frames.";
RL Yeast 11:1421-1427(1995).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; X85807; CAA59810.1; -
CC EMBL; Z72938; CAA97167.1; -
CC PIR; S60443; S60443.
CC Germline; 141465; -.
CC DR SGD; S0003385; TOS10.
CC KW Hypothetical protein.
CC SEQUENCE 217 AA; 25213 MW; E507C7EBC095CED0 CRC64;
CC
CC
CC Query Match 4.2%; Score 6; DB 1; Length 217;
CC Best Local Similarity 100.0%; Pred. No. 1.1e+02;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 60 CFEKE 65
CC Db 160 CFEKE 165
CC
CC
CC RESULT 45
CC YMS1 ARCFU
CC ID YMS1 ARCFU STANDARD; PRT; 229 AA.
CC AC Q28033;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Hypothetical protein AF2251.
CC GN AF2251.
CC OS Archaeoglobus fulgidus.
CC OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
CC OC Archaeoglobaceae; Archaeoglobus.
CC OX NCBI_TaxID=2234;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
CC RX MEDLINE=98049343; PubMed=9389475;
CC RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
CC RA Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,
CC RA Richardson D.L., Kurlavsky A.R., Graham D.E., Kyrpides N.C.,
CC RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
CC RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
CC RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
CC RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
CC RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,

```

```

RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RL reducing archaeson Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -!- SIMILARITY: Belongs to the cinA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; AE000949; AAB89006.1; -
CC PIR; C69531; C69531.
CC TIGR; AF2251; -.
CC HAMAP; MF_00226; -; 1.
CC InterPro; IPR001453; MOCF_biosynth.
CC Pfam; PF00994; MOCF_biosynth; 1.
CC ProDom; PD002460; MOCF_biosynth; 1.
CC DR TIGRFAMS; TIGR00177; molyb_syn; 1.
CC KW Hypothetical protein; Complete proteome.
CC SEQUENCE 229 AA; 25521 MW; DPF623780665EC81 CRC64;
CC
CC
CC Query Match 4.2%; Score 6; DB 1; Length 229;
CC Best Local Similarity 100.0%; Pred. No. 1.2e+02;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 91 KQFEEL 96
CC Db 219 KQFEEL 224
CC
CC
CC RESULT 46
CC YG61 FUSNN
CC ID YG61 FUSNN STANDARD; PRT; 249 AA.
CC AC Q8E1E0;
CC DT 10-OCT-2003 (Rel. 42, Created)
CC DT 10-OCT-2003 (Rel. 42, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Hypothetical UPF0082 protein FNI661.
CC GN FNI661.
CC OS Fusobacterium nucleatum (subsp. nucleatum).
CC OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
CC OC Fusobacterium.
CC OX NCBI_TaxID=76856;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=ATCC 25586;
CC RX MEDLINE=21886394; PubMed=11899109;
CC RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
CC RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
CC RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
CC RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
CC RA Fongstein M., Kyrpides N., Overbeek R.;
CC RT "Genome sequence and analysis of the oral bacterium Fusobacterium
CC nucleatum strain ATCC 25586.";
CC RL J. Bacteriol. 184:2005-2018(2002).
CC -!- SIMILARITY: Belongs to the UPF0082 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; AE010472; AAL93776.1; -
CC HAMAP; MF_00693; -; 1.

```

```

DR InterPro: IPR002876; DUF28.
DR Pfam: PF01709; DUF28; 1.
DR ProDom: PD004323; DUF28; 1.
DR TIGRFAMs: TIGR01033; TIGR01033; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 249 AA; 27768 MW; FC721F033467208A CRC64;

Query Match 4.2%; Score 6; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 ETAKKV 131
DQ |||||
DB 213 ETAKKV 218

RESULT 47
PUR7_SYNXPX STANDARD; PRT; 250 AA.
AC Q7U8Q9;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC 6.3.2.6)
DE (SAICAR synthetase).
DE (SAICAR synthetase).
GN PURC OR SYNW0554.
OS Synechococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641;
RA Palenik B., Brahamsha B., Larimer F.W., Land M., Hauser L., Chain P.,
RA Lamerdin J., Regala W., Allen E.E., McCarren J., Paulsen I.,
RA Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synechococcus."
RL Nature 424:1037-1042(2003).
CC -!- Catalytic Activity: ATP + 5-amino-1-(5-phospho-D-
CC ribosyl)imidazole-4-carboxylate + L-aspartate = ADP + phosphate +
CC (S)-2-[5-amino-1-(5-phospho-D-ribosyl)imidazole-4-
CC carboxamido]succinate.
CC -!- PATHWAY: De novo purine biosynthesis; seventh step.
CC -!- SIMILARITY: Belongs to the SAICAR synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: BX569690; CAE07069.1; -
CC HAMAP: MF 00137; -; 1.
CC InterPro: IPR001636; SAICAR synt.
CC Pfam: PF01259; SAICAR synt; 1.
CC ProDom: PD003043; SAICAR synt; 1.
CC TIGRFAMs: TIGR00081; purC; 1.
CC PROSITE: PS01057; SAICAR SYNTHETASE 1; 1.
CC PROSITE: PS01058; SAICAR SYNTHETASE 2; FALSE NEG.
KW Purine biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 250 AA; 28002 MW; 8E2B604EF31CBAF2 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 AIFQLA 139
DQ |||||
DB 152 AIFQLA 157

RESULT 48
LIVG_ARCFU STANDARD; PRT; 257 AA.
AC O28881;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable branched-chain amino acid transport ATP-binding protein livg.
DE LIVG OR AF1390.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;

```

```

OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Krelage A.S., Graham D.E., Kyripides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeill L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spriggs T., Artchik P., Kane B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.J., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -!- FUNCTION: Probable component of a branched-chain amino-acid
CC transport system.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC -----
CC EMBL; AF001008; AAB89855.1; ALT_INIT.
CC TIGR; AF1390; -.
CC InterPro; IPR003593; AAA ATPase.
CC InterPro; IPR003439; ABC transporter.
CC Pfam; PF00005; ABC tran; 1.
CC ProDom; PD000006; ABC transporter; 1.
CC SMART; SM00382; AAA; 1.
CC PROSITE; PS00811; ABC TRANSPORTER_1; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
CC Amino-acid transport; Transport; ATP-binding; Complete proteome.
KW NP BIND 36 43 ATP (POTENTIAL).
FT SEQUENCE 257 AA; 28854 MW; A8D8397E0F3B8491 CRC64;
SQ
Query Match 4.2%; Score 6; DB 1; Length 257;
Best Local Similarity 100.0%; Pred.No.1.3e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
QY 100 EFLKLD 105
DB 134 EFLKLD 139
|||||
SPRC_CABEL STANDARD; PRT; 264 AA.
ID _SPRC_CABEL
AC P34714;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SPARC precursor (Secreted protein acidic and rich in cysteine)
DE (Osteonectin) (ON) (Basement membrane protein BM-40).
DE OST-1 OR SPARC OR C4B12.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94080020; PubMed=8257796;

```

```

RA Schwarzbauer J.E., Spencer C.S.;
RT "The Caenorhabditis elegans homologue of the extracellular calcium
RT binding protein SPARC/osteonectin affects nematode body morphology
RT and mobility."
RL Mol. Biol. Cell 4:941-952(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Tin-Wollam A.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HAS A HIGH AFFINITY FOR COLLAGEN. AFFECTS NEMATODE
CC BODY MORPHOLOGY AND MOBILITY. ESSENTIAL FOR C.ELEGANS DEVELOPMENT
CC AND MUSCLE FUNCTION. THE CYSTEINE-RICH REGION COULD HAVE PROTEASE
CC INHIBITORY ACTIVITY OR MAY PROVIDE THE FRAMEWORK FOR A PROTEIN
CC BINDING MODULE. PROBABLE ROLE IN SKELETAL MORPHOGENESIS.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- TISSUE SPECIFICITY: EXPRESSED BY BODY WALL AND SEX MUSCLE CELLS.
CC PROBABLE ASSOCIATION WITH BASEMENT MEMBRANES.
CC -!- DEVELOPMENTAL STAGE: FIRST EXPRESSED IN UNHATCHED LARVAE;
CC CONTINUOUS THROUGHOUT SUBSEQUENT LARVAL STAGES AND IN ADULTS.
CC -!- SIMILARITY: Contains 1 osteonectin-like domain.
CC -!- SIMILARITY: Contains 1 Kazal-like domain.
CC -!- SIMILARITY: Contains 1 EF-hand calcium-binding domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC -----
CC EMBL; L21758; AAA16827.1; -.
CC EMBL; AF036692; AAB88325.1; -.
CC PIR; A47737; A47737.
CC HSP; P09486; ISRA.
CC WormPep; C44B12.2; CE08703.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR003645; FcIIN.
CC InterPro; IPR002350; Kazal.
CC InterPro; IPR001999; Osteonectin.
CC Pfam; PF00050; Kazal; 1.
CC SMART; SM00274; FcIIN; 1.
CC SMART; SM00280; KAZAL; 1.
CC PROSITE; PS00018; EF-HAND; 1.
CC PROSITE; PS00612; OSTEONECTIN_1; 1.
CC PROSITE; PS00613; OSTEONECTIN_2; 1.
CC Extracellular matrix; Glycoprotein; Calcium-binding; Copper; Signal;
KW Developmental protein; Basement membrane.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 264 SPARC.
FT DOMAIN 75 135 KAZAL-LIKE.
FT CA_BIND 237 248 EF-HAND (BY SIMILARITY).
FT DISULFID 53 64 BY SIMILARITY.
FT DISULFID 58 74 BY SIMILARITY.
FT DISULFID 76 110 BY SIMILARITY.
FT DISULFID 80 103 BY SIMILARITY.
FT DISULFID 92 135 BY SIMILARITY.
FT DISULFID 141 228 BY SIMILARITY.
FT DISULFID 236 252 BY SIMILARITY.
FT CARBOHYD 96 96 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 264 AA; 30172 MW; 2P952056B466E18F CRC64;
Query Match 4.2%; Score 6; DB 1; Length 264;
Best Local Similarity 100.0%; Pred.No.1.4e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
QY 26 PFLEGC 31
DB 231 PFLEGC 236
|||||

```

```

RESULT 51
TRY6_ANOGA
ID TRY6_ANOGA STANDARD; PRT; 273 AA.
AC P35040;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Trypsin 6 precursor (EC 3.4.21.4).
GN TRYP6.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Suakoko; TISSUE=Midgut;
RX MEDLINE=9332779; PubMed=8335004;
RA Mueller H.M., Crampton J.M., della Torre A., Sinden R., Crisanti A.;
RT "Members of a trypsin gene family in Anopheles gambiae are induced in
RT the gut by blood meal."
RL EMBO J. 12:2891-2900(1993).
CC -!- FUNCTION: Major function may be to aid in digestion of the blood
CC meal.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Midgut.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- CAUTION: As this protein has Gly-222 in the position that
CC determines the specificity of the enzyme instead of the Asp found
CC in trypsins, it could have a chymotrypsin-like activity.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z22930; CAA80513.1; -
CC DR PIR: S40003; S40003.
CC DR HSP: P00761; IEPT.
CC DR MEROPS: S01.112; -.
CC DR InterPro: IPR009003; Cys Ser trypsin.
CC DR InterPro: IPR001254; Peptidase S1.
CC DR InterPro: IPR001314; Peptidase_S1A.
CC DR Pfam: PF00089; trypsin; 1.
CC DR PRINTS: PR00722; CHYMOTRYPSIN.
CC DR SMART: SM00020; TRYP_SPC; 1.
CC DR PROSITE: PSS00240; TRYP_SIN_DOM; 1.
CC DR PROSITE: PS00134; TRYP_SIN_HIS; 1.
CC DR PROSITE: PS00135; TRYP_SIN_SER; FALSE NEG.
CC KW Hydrolyase; Serine protease; Signal; Zymogen; Digestion;
CC Multigene family.
CC FT SIGNAL 1 22 POTENTIAL.
CC FT PROPEP 23 46 ACTIVATION PEPTIDE.
CC FT CHAIN 47 273 TRYPSIN 6.
CC FT ACT_SITE 87 87 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 132 132 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 228 228 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT DISULFID 72 88 BY SIMILARITY.
CC FT DISULFID 197 213 BY SIMILARITY.
CC FT DISULFID 224 248 BY SIMILARITY.
CC FT SITE 222 222 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
CC SQ SEQUENCE 273 AA; 29373 MW; 83ED22152BFD2D66 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 DDPICE 76
Db 157 DDPICE 162

```

```

RESULT 52
SPEE_PVRFU
ID SPEE_PVRFU STANDARD; PRT; 281 AA.
AC Q8U4G1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable spermidine synthase (EC 2.5.1.16) (Putrescine
DE aminopropyltransferase) (SPDSV).
GN SPEE OR PF0127.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb P.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: S-adenosylmethionine + putrescine = 5'-
CC methylthioadenosine + spermidine.
CC -!- PATHWAY: Spermidine biosynthesis; fifth (last) step.
CC -!- SIMILARITY: Belongs to the spermidine/spermine synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB010139; AAL80251.1; -
CC DR PDB: 1MJF; 18-SEP-02.
CC DR HAMAP: MF_00198; -; 1.
CC DR InterPro: IPR000051; SAM_bind.
CC DR InterPro: IPR001045; Spermine synthase.
CC DR Pfam: PF01564; Spermine synth; 1.
CC DR TIGRFAMs: TIGR00417; speE; 1.
CC DR PROSITE: PS01330; SPERMIDINE SYNTHASE; 1.
CC KW Spermidine biosynthesis; Transferase; Complete proteome; 3D-structure.
CC FT DOMAIN 77 117 BINDING TO DECARBOXYLATED SAM
CC FT POTENTIAL.
CC SQ SEQUENCE 281 AA; 32334 MW; 72C97AE45BE1C487 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 DRERAK 110
Db 247 DRERAK 252

RESULT 53
Y765_HAEIN
ID Y765_HAEIN STANDARD; PRT; 282 AA.
AC Q57125; C05033;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical glycosyl transferase HI0765 (EC 2.---).
GN HI0765.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;

```

```

RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerkvliet A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 25.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32760; AAC22423.1; -.
CC PIR: F64091; F64091.
CC TIGR: H10765; -.
CC InterPro: IPR002654; Glyco_transf_25; 1.
CC Pfam: PF01755; Glyco_transf_25; 1.
CC Hypothetical protein; Transferase; Glycosyltransferase;
KW Complete proteome.
SQ SEQUENCE 282 AA; 32481 MW; D962E402AA207A18 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 282;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 BETAXK 130
DB 72 BETAXK 77

RESULT 54
FPG_STRCO
ID_FPG_STRCO STANDARD; PRT; 286 AA.
AC Q9ZBQ6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23) (FAPY-DNA
DE glycosylase).
GN MUTM OR FPG OR SC05573 OR SC7A1.17.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Batenan A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinovitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wierzbicki A., Woodward J., Barrall B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- FUNCTION: This enzyme may play a significant role in processes
CC leading to recovery from mutagenesis and/or cell death by

```

```

CC alkylating agents (By similarity).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of DNA containing ring-opened N(7)-
CC methylguanine residues, releasing 2,6-diamino-4-hydroxy-5-(N-
CC methyl)formamidopyrimidine.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SIMILARITY: Belongs to the FPG family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ939124; CAA22416.1; -.
CC PIR: T35657; T35657.
CC HSP: OS0606; 1EES.
CC HAMAP: MF_00103; -.
CC InterPro: IPR000191; Fapy DNA glyco.
CC InterPro: IPR000214; Fapy_DNAG_Zn_BS.
CC Pfam: PF01149; Fapy_DNA_glyco; 1.
CC ProDom: PD003680; Fapy_DNA_glyco; 1.
CC TIGRFAMS: TIGR00577; fpg; 1.
CC PROSITE: PS01242; FPG; 1.
CC DNA repair; Hydrolase; Glycosidase; Zinc; Zinc-finger;
KW Complete proteome. 278 POTENTIAL.
FT ZN FING 255
SQ SEQUENCE 286 AA; 32511 MW; C98F08045A9F386B CRC64;

Query Match 4.2%; Score 6; DB 1; Length 286;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LKDHRI 19
DB 48 LKDHRI 53

RESULT 55
SGBU_HAEIN
ID_SGBU_HAEIN STANDARD; PRT; 286 AA.
AC P44990;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative hexulose-6-phosphate isomerase (EC 5.-.-.-) (HUMPI).
GN SGBU OR H1026.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerkvliet A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: Isomerization of D-arabino-6-hexulose 3-phosphate to D-
CC fructose 6-phosphate.
CC -!- PATHWAY: Probably part of a sugar metabolic pathway along with
CC sgbH.
CC -!- SIMILARITY: BELONGS TO THE HUMPI FAMILY.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U32783; AAC22686.1; -  
 DR PIR; G64164; G64164.  
 DR TIGR; H1026; -

DR InterPro; IPR004560; Hx16Piso\_put.  
 DR Pfam; PF03809; Hx16Piso\_put; 1.  
 DR TIGRFAMs; TIGR00542; hx16Piso\_put; 1.  
 DR isomerase; Complete proteome.  
 KW SEQUENCE 286 AA; 33062 MW; 689543EBDE5E89BB CRC64;  
 SQ

Query Match 4.2%; Score 6; DB 1; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 BELTLG 99  
 DB 196 BELTLG 201

RESULT 56  
 NOLT\_RHIFR  
 ID NOLT\_RHIFR STANDARD; PRT; 289 AA.  
 AC P33209;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Nodulation protein NOLT precursor.  
 GN NOLT  
 OS Rhizobium fredii (Sinorhizobium fredii).  
 OG Plasmid sym.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
 OX NCBI\_TaxID=380;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=USDA 257;  
 RX MEDLINE=94018604; PubMed=8412662;  
 RA Meinhardt L.W., Krishnan H.B., Balatti P.A., Pueppke S.G.;  
 RT "Molecular cloning and characterization of a sym plasmid locus that  
 RT regulates cultivar-specific nodulation of soybean by Rhizobium fredii  
 RT USDA257".  
 RL Mol. Microbiol. 9:17-29(1993).  
 CC -!- FUNCTION: Regulates cultivar-specific nodulation of soybean.  
 CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid  
 CC anchor (Probable).  
 CC -!- INDUCTION: By flavonoid signal compounds.  
 CC -!- SIMILARITY: Belongs to the yscJ lipoprotein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L12251; AAB17677.1; -  
 DR PIR; S35022; S35022.  
 DR InterPro; IPR000437; Prok\_lipoprot\_S.  
 DR InterPro; IPR003282; SecIIOMP\_K.  
 DR InterPro; IPR006182; YscJ\_Flip.  
 DR Pfam; PF01514; YscJ\_Flip; 1.  
 DR PRINTS; PR01338; TYE3OMKPROT.  
 DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; FALSE NEG.  
 KW Plasmid; Outer membrane; Signal; Lipoprotein; Nodulation;

KW Transmembrane; Palmitate.  
 FT SIGNAL 1 33 PROBABLE.  
 FT CHAIN 34 289 MODULATION PROTEIN NOLT.  
 FT LIPID 34 34 N-palmitoyl cysteine (Potential).  
 FT LIPID 34 34 S-diacylglycerol cysteine (Potential).  
 FT TRANSMEM 233 253 POTENTIAL.  
 SQ SEQUENCE 289 AA; 31241 MW; 5944DE5C9A60DB2B CRC64;

Query Match 4.2%; Score 6; DB 1; Length 289;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EPDLAQ 56  
 DB 216 EPDLAQ 221

## RESULT 57

CHIA\_CUCSA  
 ID CHIA\_CUCSA STANDARD; PRT; 292 AA.  
 AC P17541;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Acidic endochitinase precursor (EC 3.2.1.14).  
 OS Cucumis sativus (Cucumber).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.  
 OX NCBI\_TaxID=3659;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
 RC STRAIN=cv Wisconsin SNR-58;  
 RX MEDLINE=89128884; PubMed=2915985;  
 RA Mettraux J.P., Burkhardt W., Moyer M., Dancher S., Middlesteadt W.,  
 RA Williams S., Payne G., Carnes M., Ryals J.;  
 RT "Isolation of a complementary DNA encoding a chitinase with  
 RT structural homology to a bifunctional lysozyme/chitinase".  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:896-900(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Leaf;  
 RX MEDLINE=94220741; PubMed=8167370;  
 RA Lawton K., Beck J., Potter S., Ward E., Ryals J.;  
 RT "Regulation of cucumber class III chitinase gene expression.";  
 RL Mol. Plant Microbe Interact. 7:48-57(1994).  
 CC -!- FUNCTION: This protein functions as a defense against chitin  
 CC containing fungal pathogens.  
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-  
 CC acetyl-D-glucosamine polymers of chitin.  
 CC -!- SUBCELLULAR LOCATION: Extracellular.  
 CC -!- INDUCTION: By salicylate and upon tobacco necrosis virus  
 CC infection.  
 CC -!- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl  
 CC hydrolases).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M24365; AAC33120.1; -  
 DR EMBL; M84214; AAC37395.1; -  
 DR PIR; A31455; A31455.  
 DR HGSP; P23472; 2HVM  
 DR InterPro; IPR001223; Glyco\_hydro\_18.  
 DR InterPro; IPR001579; Glyco\_hydro\_18AS.  
 DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
 DR PROSITE; PS01095; CHITINASE\_18; 1.  
 KW Hydrolyase; Glycosidase; Chitin degradation; Signal.



```

FT SIGNAL 1 25
FT CHAIN 26 292
FT DISULFID 45 92
FT DISULFID 75 82
FT DISULFID 180 209
FT ACT SITE 152 152
SQ SEQUENCE 292 AA; 30774 MW; 2DE39D42BDBE0093 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 292;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 GCACFLS 88
Db 81 GCACFLS 86

RESULT 58
ENGC_BACAA STANDARD; PRT; 293 AA.
AC Q8IWH7;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Probable GTPase engC (EC 3.6.1.-).
GN ENGC OR BA3999
OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=198094;
RN [1]
RX MEDLINE=22608414; PubMed=12721629;
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapfel E.K., Ostad O.A., Helgason E., Ralston J., Wu M.,
RA Kolonay J.F., Beaman M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Barry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R., Redmond C., Thwaiter J.E., White O., Salzberg S.L.,
RA Thomson B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.B.,
RA Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
CC -1- FUNCTION: Unusual circularly permuted GTPase that catalyzes rapid
CC hydrolysis of GTP with a slow catalytic turnover (By similarity).
CC -1- SUBUNIT: Monomer (Probable).
CC -1- SIMILARITY: Contains 1 engC GTPase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; AE017036; AAP27727.1; -.
CC DR TIGR; BA3999; -.
CC DR HAMAP; MF_01820; -.
CC DR InterPro; IPR004881; DUF258.
CC DR Pfam; PF03193; DUF258; 1.
CC DR TIGRFAMs; TIGR00157; TIGR00157; 1.
CC DR PROSITE; PS50936; ENGC_GTPASE; 1.
CC DR Hydrolase; GTP-binding; Complete proteome.
FT DONAIN 72 221
FT NP BIND 112 115
FT NP BIND 166 173
FT NP BIND 216 220
FT SITE 247 260
SQ SEQUENCE 293 AA; 33132 MW; 315FD5AF9645E875 CRC64;

```

```

Query Match 4.2%; Score 6; DB 1; Length 293;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 PFLEGC 31
Db 155 PFLEGC 160

RESULT 59
Y237_MYCPN STANDARD; PRT; 294 AA.
AC F75455;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG237 homolog (F10_orf294).
GN MPR330 OR MP506.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RX STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Piagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; AE000050; AAB96154.1; -.
CC DR PIR; S73832; S73832.
CC DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 294 AA; 34135 MW; DB9F1896E24073AC CRC64;

Query Match 4.2%; Score 6; DB 1; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 LTLGEF 101
Db 190 LTLGEF 195

RESULT 60
Y680_HAEIN STANDARD; PRT; 298 AA.
AC Q57389; O05030;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein HI0680.
GN HI0680.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RX STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

```

```
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodok A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudex D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.,
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.",
RL Science 269:496-512(1995).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -!- SIMILARITY: Belongs to the eamA transporter family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U32750; AAC22339.1; -.
CC PIR: I64085; I64085.
CC TIGR: H10680; -.
CC InterPro: IPR000620; DUF6.
CC InterPro: IPR004626; RarD.
CC Pfam: PF00892; DUF6; 1.
CC TIGRFAMs: TIGR00688; rarD; 1.
CC Hypothetical protein; Transport; Transmembrane; Complete proteome.
KW
FT TRANSMEM 5 23
FT TRANSMEM 33 52
FT TRANSMEM 72 91
FT TRANSMEM 101 120
FT TRANSMEM 127 145
FT TRANSMEM 149 166
FT TRANSMEM 175 194
FT TRANSMEM 207 229
FT TRANSMEM 238 260
FT TRANSMEM 265 284
FT TRANSMEM 298 346; 33463 MW; 65DC2A9EB8F9EF02 CRC64;
SQ SEQUENCE 298 AA; 33463 MW; 65DC2A9EB8F9EF02 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 RISTFK 23
DB 123 RISTFK 128

RESULT 61
LXD2 PHOLE STANDARD; PRT; 305 AA.
AC Q06878;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Acyl transferase (EC 2.3.1.-) (ACT) (Myristoyl-ACP-specific
DE thioesterase).
DE LUXD.
GN LUXD.
OS Photobacterium leiognathi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=658;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=741;
RA MEDLINE=93321531; PubMed=8472957;
RX Chao Y.-F., Weng S.-F., Lin J.-W.;
RT "Sequence of the luxD gene encoding acyltransferase of the lux operon
RT from Photobacterium leiognathi."
RL Gene 126:155-156(1993).
CC -!- FUNCTION: ACYL TRANSFERASE IS PART OF THE FATTY ACID REDUCTASE
```

```
CC SYSTEM REQUIRED FOR ALDEHYDE BIOSYNTHESIS; IT PRODUCES FATTY
CC ACIDS FOR THE LUMINESCENT REACTION.
CC -!- PATHWAY: Bioluminescent fatty acid reduction system;; first step.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X65612; CAA46563.1; -.
CC PIR: JN0518; JN0518.
CC HSP: P05531; 1THT.
CC InterPro: IPR003157; Acyl transf.
CC Pfam: PF02773; Acyl_transf_2; 1.
CC Luminescence; Transferase; Acyltransferase.
CC SEQUENCE 305 AA; 34384 MW; 0BDAA76D08913164 CRC64;
SQ SEQUENCE 305 AA; 34384 MW; 0BDAA76D08913164 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 305;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 KNKIAK 115
DB 296 KNKIAK 301

RESULT 62
OTCC MYCCC STANDARD; PRT; 312 AA.
AC P59779;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ornithine carbamoyltransferase, catabolic (EC 2.1.3.3) (OTCase).
DE ARCB.
OS Mycoplasma capricolum subsp. capripneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=40480;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GL100;
RA Freyraud A., Thiaucourt F.;
RT "A specific PCR for Mycoplasma putrefaciens.";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
CC + L-citrulline.
CC -!- PATHWAY: Arginine degradation; arginine deiminase pathway; second
CC step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the ArgCase/OTCase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AY282502; AAP38180.1; -.
CC HAMAP: MF_01109; -.
CC InterPro: IPR006130; Asp/Orn Cotransf.
CC InterPro: IPR002922; Orn carbtransf.
CC InterPro: IPR006131; OTCase_O.
CC InterPro: IPR006132; OTCase_P.
CC Pfam: PF00185; OTCase; 1.
CC Pfam: PF02729; OTCase_N; 1.
CC PRINTS: PR00100; AOTCASE.
CC TIGRFAMs: TIGR00658; orn carb tr; 1.
CC PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
```

KW Arginine metabolism; Transferase.  
FT SITE 32 32 IMPORTANT FOR STRUCTURAL INTEGRITY (BY  
FT SITE 57 61 CARBAMOYLPHOSPHATE BINDING (BY  
FT SITE 108 108 CARBAMOYLPHOSPHATE BINDING (BY  
FT SITE 135 135 CARBAMOYLPHOSPHATE BINDING (BY  
FT SITE 148 148 CARBAMOYLPHOSPHATE BINDING (BY  
FT SITE 271 274 ORNITHINE BINDING (BY SIMILARITY).  
FT SITE 312 AA; 35026 MW; 3761P407593F4DDB CRC64;  
SQ SEQUENCE 312 AA; 35026 MW; 3761P407593F4DDB CRC64;  
Query Match 4.2%; Score 6; DB 1; Length 312;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 113 IAKETN 118  
DB 204 IAKETN 209  
RESULT 63  
PANE\_LACLA STANDARD; PRT; 312 AA.  
AC Q9CFI8; 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Putative 2-dehydropantoate 2-reductase (EC 1.1.1.159) (Ketopantoate  
DE reductase) (KPA reductase) (KPR).  
GN LL1323.  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
OX NCBI\_TaxID=1360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21235186; PubMed=11337471;  
RA Boletín A.; Wincker P.; Mauger S.; Jaillon O.; Malarne K.,  
RA Weissbach J., Ehrlich S.D., Sorokin A.;  
RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
RT lactis ssp. lactis IL1403."  
RL Genome Res. 11:731-753(2001).  
CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of ketopantoate  
CC into pantoic acid (By similarity).  
CC -!- CATALYTIC ACTIVITY: (R)-pantoate + NADP(+) = 2-dehydropantoate +  
CC NADPH.  
CC -!- PATHWAY: Pantothenate biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -!- SIMILARITY: Belongs to the ketopantoate reductase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE006364; AAK05421.1; -.  
DR PIR; C86790; C86790.  
DR InterPro; IPR008927; 6DGDH\_C-like.  
DR InterPro; IPR003710; Apba\_.  
DR Pfam; PF02558; Apba; 1.  
DR TIGRfams; TIGR00745; apba\_pane; 1.  
KW Hypothetical protein; Pantothenate biosynthesis; Oxidoreductase; NADP;  
Complete proteome.  
FT NP\_BIND 7 12 NADP (POTENTIAL).  
FT ACT\_SITE 187 187 BY SIMILARITY.  
FT ACT\_SITE 273 273 BY SIMILARITY.  
SQ SEQUENCE 312 AA; 34443 MW; 1C7FC4547F081BAD CRC64;

KW Arginine metabolism; Transferase.  
FT SITE 32 32 IMPORTANT FOR STRUCTURAL INTEGRITY (BY  
FT SITE 57 61 CARBAMOYLPHOSPHATE BINDING (BY  
FT SITE 108 108 CARBAMOYLPHOSPHATE BINDING (BY  
FT SITE 135 135 CARBAMOYLPHOSPHATE BINDING (BY  
FT SITE 148 148 CARBAMOYLPHOSPHATE BINDING (BY  
FT SITE 271 274 ORNITHINE BINDING (BY SIMILARITY).  
FT SITE 312 AA; 35026 MW; 3761P407593F4DDB CRC64;  
SQ SEQUENCE 312 AA; 35026 MW; 3761P407593F4DDB CRC64;  
Query Match 4.2%; Score 6; DB 1; Length 312;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 113 IAKETN 118  
DB 204 IAKETN 209  
RESULT 63  
PANE\_LACLA STANDARD; PRT; 312 AA.  
AC Q9CFI8; 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Putative 2-dehydropantoate 2-reductase (EC 1.1.1.159) (Ketopantoate  
DE reductase) (KPA reductase) (KPR).  
GN LL1323.  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
OX NCBI\_TaxID=1360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21235186; PubMed=11337471;  
RA Boletín A.; Wincker P.; Mauger S.; Jaillon O.; Malarne K.,  
RA Weissbach J., Ehrlich S.D., Sorokin A.;  
RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
RT lactis ssp. lactis IL1403."  
RL Genome Res. 11:731-753(2001).  
CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of ketopantoate  
CC into pantoic acid (By similarity).  
CC -!- CATALYTIC ACTIVITY: (R)-pantoate + NADP(+) = 2-dehydropantoate +  
CC NADPH.  
CC -!- PATHWAY: Pantothenate biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -!- SIMILARITY: Belongs to the ketopantoate reductase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE006364; AAK05421.1; -.  
DR PIR; C86790; C86790.  
DR InterPro; IPR008927; 6DGDH\_C-like.  
DR InterPro; IPR003710; Apba\_.  
DR Pfam; PF02558; Apba; 1.  
DR TIGRfams; TIGR00745; apba\_pane; 1.  
KW Hypothetical protein; Pantothenate biosynthesis; Oxidoreductase; NADP;  
Complete proteome.  
FT NP\_BIND 7 12 NADP (POTENTIAL).  
FT ACT\_SITE 187 187 BY SIMILARITY.  
FT ACT\_SITE 273 273 BY SIMILARITY.  
SQ SEQUENCE 312 AA; 34443 MW; 1C7FC4547F081BAD CRC64;

Query Match 4.2%; Score 6; DB 1; Length 312;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 125 BETAKK 130  
DB 157 BETAKK 162  
RESULT 64  
CPKA\_PYRAB STANDARD; PRT; 314 AA.  
ID CPKA\_PYRAB  
AC Q9UNF8; 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Carbamate kinase (EC 2.7.2.2) (Carbamate kinase-like  
DE carbamoylphosphate synthetase).  
GN CPKA OR CPA OR PYRAB08820 OR PAB0593.  
OS Pyrococcus abyssi.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcaceae; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=29292;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=GES / Orsay;  
RX MEDLINE=21414744; PubMed=11533892;  
RA Pucarea C.; Herve G.; Cunin R.; Evans D.R.;  
RA "Cloning, expression, and structure analysis of carbamate kinase-like  
RT carbamoyl phosphate synthetase from Pyrococcus abyssi."  
RL Extremophiles 5:229-239(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=GES / Orsay;  
RX MEDLINE=22511545; PubMed=12622808;  
RA Cohen G.N.; Barbe V.; Plament D.; Galperin M.; Heilig R.; Lecompte O.,  
RA Poch O.; Prieur D.; Querellou J.; Ripp R.; Thierry J.-C.,  
RA Van der Oost J.; Weissbach J.; Zivanovic Y.; Forterre P.;  
RT "An integrated analysis of the genome of the hyperthermophilic  
RT archaean Pyrococcus abyssi."  
RL Mol. Microbiol. 47:1495-1512(2003).  
CC -!- CATALYTIC ACTIVITY: ATP + NH(3) + CO(2) = ADP + carbamoyl  
CC phosphate.  
CC -!- SUBUNIT: Homodimer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the carbamate kinase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF081680; AAF23075.1; -.  
DR EMBL; AJ248285; CAB49796.1; -.  
DR PIR; C75135; C75135.  
DR HSSP; P95474; 1E19.  
DR InterPro; IPR001048; Aa\_kinase.  
DR InterPro; IPR003964; Bac\_carb\_kinase.  
DR Pfam; PF00696; aak\_kinase; 1.  
DR PRINTS; PR01469; CARBMTKINASE.  
DR TIGRfams; TIGR00746; arcC; 1.  
KW Transferase; Kinase; Complete proteome.  
SQ SEQUENCE 314 AA; 34355 MW; 85DED48E6A01A8E2 CRC64;  
Query Match 4.2%; Score 6; DB 1; Length 314;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 125 BETAKK 130

```

Db      139 BETAKK 144
|||||
RESULT 65
NIA CHLVU
ID_ NIA CHLVU STANDARD; PRT; 318 AA.
AC Q01I70;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nitrate reductase [NADH] (EC 1.7.1.1) (NR) (Fragment).
OS Chlorella vulgaris.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
OX NCBI_TaxID=3077;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=91354204; PubMed=1883330;
RX Cannons A.C., Iida N., Solomonson L.P.;
RT "Expression of a cDNA clone encoding the haem-binding domain of
RT Chlorella nitrate reductase."
RL Biochem. J. 278:203-209(1991).
CC -!- FUNCTION: Nitrate reductase is a key enzyme involved in the first
CC step of nitrate assimilation in plants, fungi and bacteria.
CC -!- CATALYTIC ACTIVITY: Nitrite + NAD(+) + H(2)O = nitrate + NADH.
CC -!- COFACTOR: Each subunit of the enzyme contains 1 equivalent of FAD,
CC heme iron, and molybdenum-pterin as prosthetic groups. The heme
CC group is called cytochrome b-557.
CC -!- SUBUNIT: Homodimer.
CC -!- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
CC N-TERMINAL DOMAIN.
CC -!- SIMILARITY: Contains 1 cytochrome b5 heme-binding domain.
CC -!- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
CC C-TERMINAL DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X56771; CAA40090.1; -.
DR PIR; S17197; S17197.
DR HSSP; P04166; 1BSM.
DR InterPro; IPR001199; CyT B5.
DR InterPro; IPR008335; Euk_Mp_oxred.
DR InterPro; IPR007110; Ig-I_Lk6.
DR InterPro; IPR005066; Mo-co_dimer.
DR InterPro; IPR000572; Oxidored_molyb.
DR Pfam; PF00173; heme_1; 1.
DR Pfam; PF03404; Mo-co_dimer; 1.
DR PRINTS; PR00363; CYTOCHROME_B5.
DR PRINTS; PR00407; EUMOPTERIN.
DR PRODOM; PD000612; CyT_B5; 1.
DR PROSITE; PS00191; CYTOCHROME_B5_1; 1.
DR PROSITE; PS0255; CYTOCHROME_B5_2; 1.
DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; PARTIAL.
KW Oxidoreductase; Flavoprotein; FAD; NAD; Heme; Molybdenum;
KW Nitrate assimilation.
FT NON TER 1
FT METAL 251 251 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT METAL 274 274 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT NON TER 318 318
SQ SEQUENCE 318 AA; 34830 NW; E60D82FE1E98292A CRC64;

Query Match 4.2%; Score 6; DB 1; Length 318;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PFLKDH 17

Db      139 BETAKK 144
|||||
RESULT 65
NIA CHLVU
ID_ NIA CHLVU STANDARD; PRT; 318 AA.
AC Q01I70;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nitrate reductase [NADH] (EC 1.7.1.1) (NR) (Fragment).
OS Chlorella vulgaris.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
OX NCBI_TaxID=3077;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=91354204; PubMed=1883330;
RX Cannons A.C., Iida N., Solomonson L.P.;
RT "Expression of a cDNA clone encoding the haem-binding domain of
RT Chlorella nitrate reductase."
RL Biochem. J. 278:203-209(1991).
CC -!- FUNCTION: Nitrate reductase is a key enzyme involved in the first
CC step of nitrate assimilation in plants, fungi and bacteria.
CC -!- CATALYTIC ACTIVITY: Nitrite + NAD(+) + H(2)O = nitrate + NADH.
CC -!- COFACTOR: Each subunit of the enzyme contains 1 equivalent of FAD,
CC heme iron, and molybdenum-pterin as prosthetic groups. The heme
CC group is called cytochrome b-557.
CC -!- SUBUNIT: Homodimer.
CC -!- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
CC N-TERMINAL DOMAIN.
CC -!- SIMILARITY: Contains 1 cytochrome b5 heme-binding domain.
CC -!- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
CC C-TERMINAL DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X56771; CAA40090.1; -.
DR PIR; S17197; S17197.
DR HSSP; P04166; 1BSM.
DR InterPro; IPR001199; CyT B5.
DR InterPro; IPR008335; Euk_Mp_oxred.
DR InterPro; IPR007110; Ig-I_Lk6.
DR InterPro; IPR005066; Mo-co_dimer.
DR InterPro; IPR000572; Oxidored_molyb.
DR Pfam; PF00173; heme_1; 1.
DR Pfam; PF03404; Mo-co_dimer; 1.
DR PRINTS; PR00363; CYTOCHROME_B5.
DR PRINTS; PR00407; EUMOPTERIN.
DR PRODOM; PD000612; CyT_B5; 1.
DR PROSITE; PS00191; CYTOCHROME_B5_1; 1.
DR PROSITE; PS0255; CYTOCHROME_B5_2; 1.
DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; PARTIAL.
KW Oxidoreductase; Flavoprotein; FAD; NAD; Heme; Molybdenum;
KW Nitrate assimilation.
FT NON TER 1
FT METAL 251 251 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT METAL 274 274 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT NON TER 318 318
SQ SEQUENCE 318 AA; 34830 NW; E60D82FE1E98292A CRC64;

Query Match 4.2%; Score 6; DB 1; Length 318;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PFLKDH 17

Db      246 PFLKDH 251
|||||
RESULT 66
K6PF OCEIH
ID_ K6PF OCEIH STANDARD; PRT; 319 AA.
AC Q8E9D6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 6-phosphofructokinase (EC 2.7.1.11) (Phosphofructokinase)
DE (Phosphohexokinase).
DE PFKA OR PFK OR OB2172.
GN Oceanobacillus thelyensis.
OS Oceanobacillus thelyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=2220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments."
RL Nucleic Acids Res. 30:3927-3935(2002).
CC -!- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
CC fructose 1,6-bisphosphate.
CC -!- PATHWAY: Key control step of glycolysis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the phosphofructokinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP004600; BAC14128.1; -.
DR HAMAP; MF_00339; 1.
DR InterPro; IPR000023; PpfFructKinase.
DR Pfam; PF00365; PFK; 1.
DR PRINTS; PR00476; PPFRTKINASE.
DR PRODOM; PD000707; PpfFructKinase; 1.
DR PROSITE; PS00433; PHOSPHOFRUCTOKINASE; 1.
KW Kinase; Transferase; Glycolysis; Complete proteome.
SQ SEQUENCE 319 AA; 34576 NW; 32DD7454310A7086 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 319;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 RAEQL 138
|||||
Db      84 RAEQL 89
|||||
RESULT 67
ODPB_BACST
ID_ ODPB_BACST STANDARD; PRT; 324 AA.
AC P21874;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pyruvate dehydrogenase E1 component, beta subunit (EC 1.2.4.1).
DE PDHB.
GN Bacillus stearothermophilus.
OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

```

```

RC STRAIN=NCA 1503;
RX MEDLINE=90345939; PubMed=2200674;
RA Hawkins C.F., Borges A., Perham R.N.;
RT "Cloning and sequence analysis of the genes encoding the alpha and
RT beta subunits of the E1 component of the pyruvate dehydrogenase
RT multienzyme complex of Bacillus stearothermophilus.";
RL Eur. J. Biochem. 191:337-346(1990).
CC -!- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
CC conversion of pyruvate to acetyl-CoA and CO(2). It contains
CC multiple copies of three enzymatic components: pyruvate
CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
CC lipoamide dehydrogenase (E3).
CC -!- CATALYTIC ACTIVITY: Pyruvate + lipoamide = S-
CC acetyldihydrolipoamide + CO(2).
CC -!- COFACTOR: Thiamine pyrophosphate.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X53560; CAA37629.1; -
CC FIR; S14230; S14230.
CC HSP; P09061; IQSO.
CC InterPro; IPR009014; Transketo_C like.
CC InterPro; IPR005476; Transketolase_C.
CC InterPro; IPR005475; Transketolase_CR.
CC Pfam; PF02779; transket_pyr; 1.
CC Pfam; PF02780; transketolase_C; 1.
CC Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate.
KW INIT MET 0
FT
SQ SEQUENCE 324 AA; 35328 MW; F7C6085E33371384 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 324;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 126 ETAKKV 131
Db 316 ETAKKV 321

RESULT 68
K6PF MYCPU STANDARD; PRT; 326 AA.
AC Q98PFG;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 6-phosphofructokinase (EC 2.7.1.11) (Phosphofructokinase)
DE (phosphohexokinase).
GN PFKA OR MYPU 6010.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;
RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galieson F.,
RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis.";
RL Nucleic Acids Res. 29:2145-2153(2001).
CC -!- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
CC fructose 1,6-bisphosphate.
CC -!- PATHWAY: Key control step of glycolysis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
```

```

CC -!- SIMILARITY: Belongs to the phosphofructokinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL445565; CAC13774.1; -
CC FIR; A99587; A99587.
CC Mypulist; MYPU_6010; -.
CC HAMAP; MF 00339; -; 1.
CC InterPro; IPR000023; Ppfrckinase.
CC Pfam; PF00365; PFK; 1.
CC PRINTS; PR00476; PPFCKTKINASE.
CC ProDom; PD000707; Ppfrckinase; 1.
CC PROSITE; PS00433; PHOSPHOFUCTOKINASE; FALSE NEG.
CC Kinase; Transferase; Glycolysis; Complete proteome.
KW KINASE; TRANSFERASE; GLYCOLYSIS; 35832 MW; 14A0F5B3F956E5 CRC64;
SQ SEQUENCE 326 AA; 35832 MW; 14A0F5B3F956E5 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 89 VKQFE 94
Db 206 VKQFE 211

RESULT 69
HUR STRAU STANDARD; PRT; 340 AA.
AC Q00923;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Hydroxyurea phosphotransferase (EC 2.7.-.-) (Hydroxyurea resistance
DE protein).
GN HUR.
OS Streptomyces aureofaciens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1894;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10762 / CCM 3239;
RX MEDLINE=92267374; PubMed=1316866;
RA Kormanec J., Farkasovsky M., Potuchkova L., Godar S.;
RT "A gene (hur) from Streptomyces aureofaciens, conferring resistance
RT to hydroxyurea, is related to genes encoding streptomycin
RT phosphotransferase.";
RL Gene 114:133-137(1992).
CC -!- FUNCTION: THIS IS A POTENTIAL PHOSPHOTRANSFERASE THAT INACTIVATES
CC HYDROXYUREA BY PHOSPHORYLATION OF THE HYDROXY GROUP IN THE
CC HYDROXYLAMINE MOIETY.
CC -!- SIMILARITY: TO AMINOGLYCOSIDE PHOSPHOTRANSFERASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M81739; AAA26767.1; -
CC FIR; JC1125; JC1125.
CC InterPro; IPR006748; APH_6_hur.
CC Pfam; PF04655; APH_6_hur; 1.
CC Transferase; Kinase; ATP-binding.
KW DOMAIN 172 184 AMINOGLYCOSIDE BINDING (BY SIMILARITY).
FT
```

FT ACT SITE 240 240 BY SIMILARITY.  
SQ SEQUENCE 340 AA; 36049 MW; 7D16C52D350BB3E9 CRC64;  
Query Match 4.2%; Score 6; DB 1; Length 340;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 104 LDREFA 109  
|||||  
Db 296 LDREFA 301  
RESULT 70  
RECA\_MYCGE  
ID RECA\_MYCGE STANDARD; PRT; 340 AA.  
AC P47581; O49512;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE RecA protein (Recombinase A).  
GN RECA OR MG339.  
OS Mycoplasma genitalium.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2097;  
EN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 33530 / G-37;  
RX MEDLINE=96026346; PubMed=7569993;  
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,  
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,  
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
RT "The minimal gene complement of Mycoplasma genitalium.";  
RL Science 270:397-403(1995).  
RN [2]  
RP SEQUENCE OF 32-124 FROM N.A.  
RC STRAIN=ATCC 33530 / G-37;  
RX MEDLINE=94075230; PubMed=8253680;  
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;  
RT "A survey of the Mycoplasma genitalium genome by using random  
RT sequencing.";  
RL J. Bacteriol. 175:7918-7930(1993).  
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of  
CC single-stranded DNA, the ATP-dependent uptake of single-stranded  
CC DNA by duplex DNA, and the ATP-dependent hybridization of  
CC homologous single-stranded DNAs. It interacts with lexA causing  
CC its activation and leading to its autocatalytic cleavage.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (BY similarity).  
CC -!- SIMILARITY: Belongs to the recA family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U39715; AAC71564.1; -;  
CC EMBL; U01704; ABA01016.1; -;  
CC PIR; E64237; E64237.  
CC HSP; P03017; 2REB.  
CC TIGR; MG339; -;  
CC HAMAP; MF\_00268; -; 1.  
CC InterPro; IPR003593; AAA\_ATPase.  
CC InterPro; IPR001553; RecA.  
CC Pfam; PF00154; recA; 1.  
CC PRINTS; PR00142; RECA.  
CC ProDom; PD000229; RecA; 1.  
CC SMART; SM00382; AAA; 1.  
CC PROSITE; PS00321; RECA\_1; 1.

DR PROSITE; PS00162; RECA\_2; 1.  
DR PROSITE; PS00163; RECA\_3; 1.  
KW DNA damage; DNA recombination; SOS response; ATP-binding; DNA-binding;  
KW Complete proteome.  
FT NP\_BIND 67 74 ATP (BY SIMILARITY).  
FT CONFLICT 32 36 AKKNS -> CKEKH (IN REF. 2).  
SQ SEQUENCE 340 AA; 37430 MW; 0957B712CD2125C0 CRC64;  
Query Match 4.2%; Score 6; DB 1; Length 340;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 110 KXKIAK 115  
|||||  
Db 250 KXKIAK 255  
RESULT 71  
VANB\_ENTFA  
ID VANB\_ENTFA STANDARD; PRT; 342 AA.  
AC Q06893; Q47821; Q57112;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Vancomycin B-type resistance protein vanB (EC 6.3.2.-) (VanB ligase)  
DE (D-alanine--D-lactate ligase).  
GN VANB OR VANB2 OR EF2294.  
OS Enterococcus faecalis (Streptococcus faecalis).  
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
OX NCBI\_TaxID=1351;  
EN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=V583 / ATCC 700802;  
RX MEDLINE=94171087; PubMed=8125347;  
RA Evers S., Reynolds P.E., Courvalin P.;  
RT "Sequence of the vanB and ddl genes encoding D-alanine:D-lactate and  
RT D-alanine:D-alanine ligases in vancomycin-resistant Enterococcus  
RT faecalis V583.";  
RL Gene 140:97-102(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=V583 / ATCC 700802;  
RX MEDLINE=96200099; PubMed=8631706;  
RA Evers S., Courvalin P.;  
RT "Regulation of vanB-type vancomycin resistance gene expression by the  
RT vanS(B)-vanR(B) two-component regulatory system in Enterococcus  
RT faecalis V583.";  
RL J. Bacteriol. 178:1302-1309(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=V583 / ATCC 700802;  
RX MEDLINE=22550857; PubMed=12663927;  
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,  
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,  
RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,  
RA Daugherty S., DeBoy R.T., Durkin S., Kolonay R., Madupu R., Nelson W.,  
RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J.J., Khouri H.,  
RA Uterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;  
RT "Role of mobile DNA in the evolution of vancomycin-resistant  
RT Enterococcus faecalis.";  
RL Science 299:2071-2074(2003).  
RN [4]  
RP SEQUENCE OF 110-305 FROM N.A.  
RC STRAIN=V583 / ATCC 700802.  
RX MEDLINE=93178969; PubMed=8440477;  
RA Evers S., Salm D.F., Courvalin P.;  
RT "The vanB gene of vancomycin-resistant Enterococcus faecalis V583 is  
RT structurally related to genes encoding D-Ala:D-Ala ligases and  
RT glycopeptide-resistance proteins Vana and Vanc.";  
RL Gene 124:143-144(1993).  
RN [5]  
RP SEQUENCE OF 103-312 FROM N.A.  
RC STRAIN=SF300;

```
OS Human herpesvirus (type 6 / strain Uganda-1102) (HHV6).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OX Betaherpesvirinae; Roseolovirus.
RN NCBI_TaxID=10370;
RP SEQUENCE FROM N.A.
RX MEDLINE=92333249; PubMed=1321206;
RA Efstathiou S., Lawrence G.L., Brown C.M., Barrell B.G.;
RT "Identification of homologues to the human cytomegalovirus US22 gene
family in human herpesvirus 6.";
RL J. Gen. Virol. 73:1661-1671(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95266321; PubMed=7747482;
RA Gompels U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J.,
RA Martin M.E., Efstathiou S., Craxton M., Macaulay H.A.;
RT "The DNA sequence of human herpesvirus-6: structure, coding content,
and genome evolution.";
RL Virology 209:29-51(1995).
CC -1- SIMILARITY: BELONGS TO THE US22 FAMILY. STRONGEST SIMILARITY TO
CC HUMAN CYTOMEGALOVIRUS UL28 PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10082; BAA00979.1; -.
DR EMBL; X83413; CAA58434.1; -.
DR PIR; JQ1651; JQ1651.
SQ SEQUENCE 342 AA; 39355 MW; 1471AE63FB0F779E CRC64;

Query Match 4.2%; Score 6; DB 1; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 LTLGFF 101
Db 57 LTLGFF 62

RESULT 73
MURB_VIECH
ID MURB_VIECH STANDARD; PRT; 347 AA.
AC Q9KV40;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UDP-N-acetylenolpyruvylglucosamine reductase (EC 1.1.1.158) (UDP-N-
DE acetylmuramate dehydrogenase).
GN MURB OR VC0318.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -1- FUNCTION: Cell wall formation (By similarity).
```

CC -!- CATALYTIC ACTIVITY: UDP-N-acetylmuramate + NADP(+) = UDP-N-  
 CC acetyl-3-O-(1-carboxyvinyl)-D-glucosamine + NADPH.  
 CC -!- COFACTOR: FAD (By similarity).  
 CC -!- PATHWAY: Peptidoglycan biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -!- SIMILARITY: Belongs to the murB family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE004120; AAF93491.1; ALT\_INIT.  
 CC PIR; D82337; D82337.  
 CC HSP; P08373; 2MR.  
 CC TIGR; VC0318; -.  
 CC HAMAP; MF\_00037; -.  
 CC InterPro; IPR003170; MurB.  
 CC InterPro; IPR006094; Oxid FAD bind.  
 CC Pfam; PF01565; FAD binding\_4; 1.  
 CC Pfam; PF02873; MurB\_C; 1.  
 CC TIGRFAMs; TIGR00179; murB; 1.  
 CC Oxidoreductase; NADP; Flavoprotein; FAD; Cell wall; Cell division;  
 CC Peptidoglycan synthesis; Complete proteome.  
 CC KW  
 CC Peptidoglycan synthesis; Complete proteome.  
 CC FT METAL 262 262 MAGNESIUM (BY SIMILARITY).  
 CC SQ SEQUENCE 347 AA; 38271 MW; C4758B7E14A31CD6 CRC64;  
 CC  
 CC Query Match 4.2%; Score 6; DB 1; Length 347;  
 CC Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 135 IEQLAA 140  
 CC |||||  
 CC Db 17 IEQLAA 22  
 CC  
 CC RESULT 74  
 CC SYFA\_STRPN  
 CC ID SYFA\_STRPN STANDARD; PRT; 348 AA.  
 CC AC Q97S36;  
 CC DT 28-FEB-2003 (Rel. 41, Created)  
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Phenylalanine-tRNA synthetase alpha chain (EC 6.1.1.20)  
 CC DE (Phenylalanine--tRNA ligase alpha chain) (PHERS).  
 CC GN PHER OR SP0579.  
 CC OS Streptococcus pneumoniae.  
 CC OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 CC OC Streptococcus.  
 CC OX NCBI\_TaxID=1313;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=ATCC BAA-334 / TIGR4;  
 CC RX MEDLINE=21357209; PubMed=11463916;  
 CC RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
 CC Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,  
 CC Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,  
 CC Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
 CC Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,  
 CC McDonald L.A., Feldblyum T.V., Anguoli S., Dickinson T., Hickey E.K.,  
 CC Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
 CC Doughterty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
 CC RT "Complete genome sequence of a virulent isolate of Streptococcus  
 CC pneumoniae".  
 CC RL Science 293:498-506(2001).  
 CC CC -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +  
 CC dihydroxyacetonephosphate + L-phenylalanyl-tRNA(Phe).  
 CC -!- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).  
 CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
 CC Pre-tRNA synthetase alpha chain subfamily 1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE007368; AAK74733.1; -.  
 CC PIR; D95067; D95067.  
 CC TIGR; SP0579; -.  
 CC HAMAP; MF\_00281; -.  
 CC InterPro; IPR004188; Phe\_tRNA\_synth\_N.  
 CC InterPro; IPR004529; PheS.  
 CC InterPro; IPR002319; tRNA-synt 2d.  
 CC InterPro; IPR006195; tRNA ligase II.  
 CC Pfam; PF02912; Phe tRNA-synt N; 1.  
 CC Pfam; PF01409; tRNA-synt 2d; 1.  
 CC TIGRFAMs; TIGR00468; PheS; 1.  
 CC PROSITE; PS50862; AA tRNA\_LIGASE II; 1.  
 CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 CC Metal-binding; Magnesium; Complete proteome.  
 CC FT METAL 262 262 MAGNESIUM (BY SIMILARITY).  
 CC SQ SEQUENCE 348 AA; 39101 MW; 7762B76B5627652D CRC64;  
 CC  
 CC Query Match 4.2%; Score 6; DB 1; Length 348;  
 CC Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 124 FEETAK 129  
 CC |||||  
 CC Db 77 FEETAK 82  
 CC  
 CC RESULT 75  
 CC EGSA\_SULTO  
 CC ID EGSA\_SULTO STANDARD; PRT; 350 AA.  
 CC AC P58460;  
 CC DT 28-FEB-2003 (Rel. 41, Created)  
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Glycerol-1-phosphate dehydrogenase [NAD(P)] (EC 1.1.1.261) (Sn-  
 CC glycerol-1-phosphate dehydrogenase) (G-1-P dehydrogenase)  
 CC DE (Enantiomeric glycerophosphate synthase).  
 CC GN EGSA OR ST0344.  
 CC OS Sulfolobus tokodaii.  
 CC OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 CC OC Sulfolobus.  
 CC OX NCBI\_TaxID=111955;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=JCM 10545 / 7;  
 CC RX MEDLINE=21456156; PubMed=11572479;  
 CC RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
 CC Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,  
 CC Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
 CC Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,  
 CC Aoki K.-I., Masuda Y., Yanagii M., Nishimura M., Yamagishi A.,  
 CC Oshima T., Kikuchi H.;  
 CC RT "Complete genome sequence of an aerobic thermoacidophilic  
 CC Crenarchaeon, Sulfolobus tokodaii strain 7".  
 CC RL DNA Res. 8:123-140(2001).  
 CC CC -!- FUNCTION: Responsible for the formation of archaea-specific  
 CC glycerophosphate backbone of phospholipids, G-1-P, from  
 CC dihydroxyacetonephosphate (DHAP) (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Sn-glycerol-1-phosphate + NAD(P)(+) =  
 CC glycerone phosphate + NAD(P)H.  
 CC -!- PATHWAY: De novo phospholipid biosynthesis.  
 CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -!- SIMILARITY: Belongs to the glycerol-1-phosphate dehydrogenase



```

CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP000982; BAB65322.1; -.
DR HAMAP; MF_00497; -. 1.
DR InterPro; IPR001670; Fe-ADH.
DR Pfam; PF00465; Fe-ADH; 1.
KW Phospholipid biosynthesis; Oxidoreductase; NADP; Complete proteome.
SQ SEQUENCE 350 AA; 37792 MW; 539461CC2F6BD310 CRC64;
Query Match 4.2%; Score 6; DB 1; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 112 KIAKET 117
Db |||||
341 KIAKET 346

```

Search completed: August 11, 2004, 14:24:15  
Job time : 15 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 14:21:17 ; Search time 39 Seconds  
(without alignments)  
1148.810 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 142

Sequence: 1 MGAPLPPAWOPFLKDHRS.....EFETAKVRRRAIEQLAAMD 142

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

```
SPREMBL.25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	74.6	142	6	Q81009
2	58	40.8	142	6	Q9GLN5
3	41	28.9	121	11	Q923F7
4	13	9.2	157	13	Q804H7
5	12	8.5	135	13	Q90ZNO
6	12	8.5	142	13	Q9DDK0
7	12	8.5	160	13	Q8JG75
8	12	8.5	160	13	Q8JG75
9	12	8.5	160	13	Q8JG75
10	9	6.3	59	13	Q90VY6
11	9	6.3	128	13	Q90WU8
12	8	5.6	142	13	Q90WU9
13	8	5.6	195	10	Q94EC5
14	8	5.6	197	16	Q82VH3
15	8	5.6	234	2	Q9F8X4
16	8	5.6	491	5	Q95YA6

17	8	5.6	706	16	P72747
18	8	5.6	875	5	O18265
19	8	5.6	1621	5	O81LL7
20	8	5.6	1822	2	Q07290
21	7	4.9	66	2	Q57137
22	7	4.9	102	16	Q8DS63
23	7	4.9	118	16	Q8PKX9
24	7	4.9	125	12	Q91CW7
25	7	4.9	130	5	O18231
26	7	4.9	130	12	Q91CX3
27	7	4.9	149	12	Q8QXX8
28	7	4.9	152	16	Q8ZJX0
29	7	4.9	152	16	Q8ZOV6
30	7	4.9	155	16	Q89DB1
31	7	4.9	171	12	Q9QBX0
32	7	4.9	175	2	O84F94
33	7	4.9	178	16	O8XKR3
34	7	4.9	182	16	O834K3
35	7	4.9	199	2	Q9JMX9
36	7	4.9	199	16	O25261
37	7	4.9	199	16	Q9ZLV1
38	7	4.9	219	16	Q8D904
39	7	4.9	224	16	O7U7M6
40	7	4.9	228	17	O58403
41	7	4.9	228	17	Q9UY11
42	7	4.9	248	9	Q37965
43	7	4.9	250	16	Q822K6
44	7	4.9	253	16	Q92TM2
45	7	4.9	255	16	Q8EBM3
46	7	4.9	265	10	O9M521
47	7	4.9	271	17	O28569
48	7	4.9	308	10	O9M526
49	7	4.9	308	10	O9M522
50	7	4.9	308	10	O9M525
51	7	4.9	308	10	O81651
52	7	4.9	308	10	O91L41
53	7	4.9	309	10	O9M520
54	7	4.9	310	10	O9SD20
55	7	4.9	312	12	O67551
56	7	4.9	313	10	O9LD14
57	7	4.9	314	10	O9LD00
58	7	4.9	343	3	O59708
59	7	4.9	353	16	Q7USB6
60	7	4.9	374	5	O9TXL0
61	7	4.9	379	16	Q8XTR2
62	7	4.9	393	16	Q91253
63	7	4.9	405	13	O8AVV0
64	7	4.9	427	10	O9SCQ5
65	7	4.9	442	5	O26582
66	7	4.9	467	16	Q89ZX8
67	7	4.9	501	16	Q8P9W0
68	7	4.9	525	16	Q97E19
69	7	4.9	546	16	Q87YV7
70	7	4.9	579	10	O9LD30
71	7	4.9	609	16	Q87FL8
72	7	4.9	625	17	Q8PZW5
73	7	4.9	671	16	O8EUJ0
74	7	4.9	673	11	O91VK6
75	7	4.9	698	16	O8ZJ11
76	7	4.9	731	16	O7UMT4
77	7	4.9	762	10	O943W8
78	7	4.9	778	16	O8RIK8
79	7	4.9	788	16	O81I26
80	7	4.9	885	16	O55683
81	7	4.9	932	10	O49680
82	7	4.9	966	16	O89PV0
83	7	4.9	1032	3	Q8WZ22
84	7	4.9	1042	16	O7VUR9
85	7	4.9	1084	16	O8XJX8
86	7	4.9	1098	5	O19149
87	7	4.9	1205	5	O9V8B7
88	7	4.9	1770	5	O813A3
89	7	4.9	1924	12	O67654

P72747	synchocyst
O18265	caenorhabdi
O81LL7	plasmodium
Q07290	streptococc
Q57137	paracoccus
Q8DS63	streptococc
Q8PKX9	xanthomonas
Q91CW7	kalanchoe 1
Q18231	caenorhabdi
Q91CX3	kalanchoe 1
Q8QXX8	garlic late
Q8ZJX0	salmonella
Q8ZOV6	salmonella
Q9QBX0	bradyrhizob
Q9QBX0	lily latent
O84F94	myxococcus
O8XKR3	clostridium
O834K3	enterococcu
Q9JMX9	helicobacte
O25261	helicobacte
Q9ZLV1	helicobacte
Q8D904	vibrio vuln
O7U7M6	synchococc
O58403	pyrococcus
Q9UY11	pyrococcus
Q37965	bacterioph
Q822K6	enterococcu
Q92TM2	rhizobium m
Q8EBM3	shewanella
O9M521	tsuga heter
O28569	archaeoglob
Q9M526	tsuga heter
Q9M522	tsuga heter
O81651	pinus taeda
O91L41	pinus taeda
O9M520	tsuga heter
O9SD20	glycine max
O67551	garlic late
O9LD14	thuja plica
O9LD00	thuja plica
O59708	schizosacch
Q7USB6	rhodopirell
O9TXL0	caenorhabdi
Q8XTR2	deinococcus
Q91253	pseudomonas
O8AVV0	xenopus lae
O9SCQ5	arabidopsis
O26582	schistosoma
Q89ZX8	bacteroides
Q8P9W0	leptospiro
Q97E19	clostridium
Q87YV7	pseudomonas
O9LD30	cryptocodi
Q87FL8	vibrio para
Q8PZW5	methanosarc
O8EUJ0	mycoplasma
O91VK6	mus muscucu
O8ZJ11	versinia pe
Q7UMT4	rhodopirell
Q943W8	oryza sativ
O8RIK8	fusobacteri
O81I26	bacillus ce
O55683	synchocyst
O49680	arabidopsis
O89PV0	bradyrhizob
Q8WZ22	neurospora
O7VUR9	helicobacte
O8XJX8	clostridium
O19149	caenorhabdi
O9V8B7	erosophila
O813A3	plasmodium
O67654	garlic late

90 Q8QY8 garlic late  
 91 Q65652 blueberry s  
 92 Q96171 plasmodium  
 93 Q81204 plasmodium  
 94 Q94657 plasmodium  
 95 Q64046 bacterioph  
 96 Q31976 bacillus su  
 97 P90580 plasmodium  
 98 Q81640 plasmodium  
 99 Q81640 plasmodium  
 100 Q97RG6 streptococc

7 4.9 1924 12 Q8QY8  
 7 4.9 1967 12 Q65652  
 7 4.9 2010 5 Q96171  
 7 4.9 2203 5 Q81204  
 7 4.9 2212 5 Q94657  
 7 4.9 2285 9 Q64046  
 7 4.9 2285 16 Q31976  
 7 4.9 2647 5 P90580  
 7 4.9 2860 5 Q81640  
 7 4.9 3384 5 Q81640  
 6 4.2 35 16 Q97RG6

## ALIGNMENTS

## RESULT 1

Q81009 PRELIMINARY; PRT; 142 AA.  
 AC Q81009;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Survivin.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Inoue C., Yamazaki J., Kano R., Hasegawa A.;  
 RT "Canis familiaris mRNA for survivin-protein, complete cds."  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB095108; BAC22748.2; -.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.  
 DR GO; GO:0006916; P:anti-apoptosis; IEA.  
 DR InterPro; IPR001370; BIR.  
 DR Pfam; PF00653; BIR; 1.  
 DR SMART; SM00238; BIR; 1.  
 DR PROSITE; PS50143; BIR\_REPEAT\_2; 1.  
 SQ SEQUENCE 142 AA; 16375 MW; E960DC9DE6D36792 CRC64;

Query Match 74.6%; Score 106; DB 6; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 3e-104;  
 Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 RMAEAGTHCPTNEPDLAQCFFCFKELGWEPPDDPIEHHKHSSGCAFLSVKKQFEEL 96  
 DB 37 RMAEAGTHCPTNEPDLAQCFFCFKELGWEPPDDPIEHHKHSSGCAFLSVKKQFEEL 96  
 QY 97 TLGFLKLDREAKNKIAETNNKKKFEETAKKVRRAIQLAAMD 142  
 DB 97 TLGFLKLDREAKNKIAETNNKKKFEETAKKVRRAIQLAAMD 142

## RESULT 2

Q9GLN5 PRELIMINARY; PRT; 142 AA.  
 AC Q9GLN5;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Apoptosis inhibitor survivin.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=21328903; PubMed=11435714;  
 RX Cirera S., Fredholm M.;  
 RT "Isolation and mapping the pig homologs survivin (BIRC5) and effector

RT cell protease receptor 1 (EPRI) genes."  
 RL CytoGenet. Cell Genet. 92:351-352 (2001).  
 DR EMBL; AF195781; AAG17540.1; -.  
 DR HSSP; O15392; IE31.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.  
 DR GO; GO:0006916; P:anti-apoptosis; IEA.  
 DR InterPro; IPR001370; BIR.  
 DR Pfam; PF00653; BIR; 1.  
 DR SMART; SM00238; BIR; 1.  
 DR PROSITE; PS50143; BIR\_REPEAT\_2; 1.  
 SQ SEQUENCE 142 AA; 16297 MW; 56BAV53E651F4F2C CRC64;

Query Match 40.8%; Score 58; DB 6; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 3e-53;  
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 AGFIHCPTNEPDLAQCFFCFKELGWEPPDDPIEHHKHSSGCAFLSVKKQFEEL 98  
 DB 41 AGFIHCPTNEPDLAQCFFCFKELGWEPPDDPIEHHKHSSGCAFLSVKKQFEEL 98

## RESULT 3

Q923F7 PRELIMINARY; PRT; 121 AA.  
 ID Q923F7;  
 AC Q923F7;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Baculoviral IAP repeat-containing 5.  
 GN BIRC5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC004702; AA04702.1; -.  
 DR MGD; MGI:1203517; Birc5.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0008189; F:apoptosis inhibitor activity; IMP.  
 DR GO; GO:0006916; P:anti-apoptosis; IMP.  
 DR InterPro; IPR001370; BIR.  
 DR Pfam; PF00653; BIR; 1.  
 DR SMART; SM00238; BIR; 1.  
 DR PROSITE; PS50143; BIR\_REPEAT\_2; 1.  
 SQ SEQUENCE 121 AA; 14154 MW; 1E0ECT7E01BA65585 CRC64;

Query Match 28.9%; Score 41; DB 11; Length 121;  
 Best Local Similarity 100.0%; Pred. No. 3e-35;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 CACTPERMAEAGTHCPTNEPDLAQCFFCFKELGWEPPDD 71  
 DB 31 CACTPERMAEAGTHCPTNEPDLAQCFFCFKELGWEPPDD 71

## RESULT 4

Q804H7 PRELIMINARY; PRT; 157 AA.  
 ID Q804H7;  
 AC Q804H7;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE SIX.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]

```

RP SEQUENCE FROM N.A.
RA Song K.-H., Kim T.-M., Kim H.-J., Kim J.W., Kim H.-H., Kwon H.-B.,
RA Kim W.S., Choi H.-S.;
RT "Molecular cloning and characterization of a novel inhibitor of
RT apoptosis protein from xenopus laevis.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY174765; AAC00085.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO; GO:0006916; P:anti-apoptosis; IEA.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 1.
DR SMART; SM00238; BIR; 1.
DR PROSITE; PS50143; BIR REPEAT 2; 1.
SQ SEQUENCE 157 AA; 18311 MW; 1ED9COA65C6ED627 CRC64;

Query Match      9.2%; Score 13; DB 13; Length 157;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 CACTPERVAEGF 43
Db 44 CACTPERVAEGF 56

RESULT 5
Q90ZNO PRELIMINARY; PRT; 135 AA.
AC Q90ZNO;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Survivin delta.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Langer J.S., Bridgham J.T., Johnson A.L.;
RT "Molecular cloning and characterization of novel chicken survivin
RT splicing variants.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF377323; AAK5308.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO; GO:0006916; P:anti-apoptosis; IEA.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 1.
DR SMART; SM00238; BIR; 1.
DR PROSITE; PS50143; BIR REPEAT 2; 1.
SQ SEQUENCE 135 AA; 14995 MW; 60D205839B465C3C CRC64;

Query Match      8.5%; Score 12; DB 13; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KELEGWEPDDDP 73
Db 64 KELEGWEPDDDP 75

RESULT 6
Q9DDKO PRELIMINARY; PRT; 142 AA.
AC Q9DDKO;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Survivin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Langer J.S., Johnson A.L., Bridgham J.T.;
RT "Molecular cloning and characterization of chicken survivin.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF322051; AAG42494.1; -.
DR HSP; O15392; IE31.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO; GO:0006916; P:anti-apoptosis; IEA.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 1.
DR SMART; SM00238; BIR; 1.
DR PROSITE; PS50143; BIR REPEAT 2; 1.
SQ SEQUENCE 142 AA; 15955 MW; 1323CCBD5ER2F551 CRC64;

Query Match      8.5%; Score 12; DB 13; Length 142;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KELEGWEPDDDP 73
Db 64 KELEGWEPDDDP 75

RESULT 7
Q8JG75 PRELIMINARY; PRT; 160 AA.
AC Q8JG75;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Survivin.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Bolton M.A., Lan W., Powers S.E., McClelland M.L., Kuang J.,
RA Stukenberg P.T.;
RT "Aurora B kinase exists in a complex with survivin and INCENP and its
RT kinase activity is stimulated by survivin binding and
RT phosphorylation.";
RL Mol. Biol. Cell 0:0-0(2002).
DR EMBL; AY115553; AAM76714.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO; GO:0006916; P:anti-apoptosis; IEA.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 1.
DR SMART; SM00238; BIR; 1.
DR PROSITE; PS50143; BIR REPEAT 2; 1.
SQ SEQUENCE 160 AA; 18677 MW; B7C32E02119DC61E CRC64;

Query Match      8.5%; Score 12; DB 13; Length 160;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KELEGWEPDDDP 73
Db 71 KELEGWEPDDDP 82

RESULT 8
Q8JGN5 PRELIMINARY; PRT; 160 AA.
AC Q8JGN5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)

```

```

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Survivin/XBIR1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Losada A., Hirano M., Hirano T.;
RT "Aurora B and polo kinases cooperate to establish metaphase chromosome
structure with bi-oriented sister chromatids."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY100639; AA044085.1;
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO: GO:0006916; P:anti-apoptosis; IEA.
DR InterPro: IPR001370; BIR.
DR Pfam: PF00653; BIR; 1.
DR SMART: SM00238; BIR; 1.
DR PROSITE: PS00143; BIR_REPEAT_2; 1.
SQ SEQUENCE 160 AA; 18686 MW; 9AC32E02119DC609 CRC64;

Query Match 8.5%; Score 12; DB 13; Length 160;
Best Local Similarity 100.0%; Pred.No. 0.00024;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KELEGWEPDDDP 73
Db 71 KELEGWEPDDDP 82

RESULT 9
Q8AXI6 PRELIMINARY; PRT; 160 AA.
ID Q8AXI6;
AC Q8AXI6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Survivin.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Murphy C.R., Sabel J.L., Sandler A.D., Dagle J.M.;
RT "Survivin mRNA is down-regulated during early Xenopus laevis
RT embryogenesis."
RL Dev. Dyn. 225:597-601(2002).
DR EMBL: AF424292; AA076690.1;
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO: GO:0006916; P:anti-apoptosis; IEA.
DR InterPro: IPR001370; BIR.
DR Pfam: PF00653; BIR; 1.
DR SMART: SM00238; BIR; 1.
DR PROSITE: PS00143; BIR_REPEAT_2; 1.
SQ SEQUENCE 160 AA; 18812 MW; 57538991C6190F4B CRC64;

Query Match 8.5%; Score 12; DB 13; Length 160;
Best Local Similarity 100.0%; Pred.No. 0.00024;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KELEGWEPDDDP 73
Db 71 KELEGWEPDDDP 82

RESULT 10
Q8AXI6 PRELIMINARY; PRT; 160 AA.
ID Q8AXI6;
AC Q8AXI6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Survivin.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Murphy C.R., Sabel J.L., Sandler A.D., Dagle J.M.;
RT "Survivin mRNA is down-regulated during early Xenopus laevis
RT embryogenesis."
RL Dev. Dyn. 225:597-601(2002).
DR EMBL: AF424292; AA076690.1;
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO: GO:0006916; P:anti-apoptosis; IEA.
DR InterPro: IPR001370; BIR.
DR Pfam: PF00653; BIR; 1.
DR SMART: SM00238; BIR; 1.
DR PROSITE: PS00143; BIR_REPEAT_2; 1.
SQ SEQUENCE 160 AA; 18812 MW; 57538991C6190F4B CRC64;

Query Match 8.5%; Score 12; DB 13; Length 160;
Best Local Similarity 100.0%; Pred.No. 0.00024;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KELEGWEPDDDP 73
Db 71 KELEGWEPDDDP 82

RESULT 11
Q90WU8 PRELIMINARY; PRT; 128 AA.
ID Q90WU8;
AC Q90WU8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Survivin 2.
GN BIR5B.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Inohara N., Nunez G.;
RT MEDLINE=20373792; PubMed=10917738;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish."
RL Cell Death Differ. 7:509-510(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Inohara N., Nunez G.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY057058; AA118251.1;
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO: GO:0006916; P:anti-apoptosis; IEA.
DR InterPro: IPR001370; BIR.
DR Pfam: PF00653; BIR; 1.
DR SMART: SM00238; BIR; 1.
DR PROSITE: PS00143; BIR_REPEAT_2; 1.
SQ SEQUENCE 128 AA; 15294 MW; AB496849082DDC66 CRC64;

Query Match 6.3%; Score 9; DB 13; Length 128;
Best Local Similarity 100.0%; Pred.No. 0.31;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 ELEGWEPDD 71
Db 52 ELEGWEPDD 60

```

```

Q90VY6 PRELIMINARY; PRT; 59 AA.
ID Q90VY6;
AC Q90VY6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Survivin gamma (Survivin short).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Langer J.S., Brigham J.T., Johnson A.L.;
RT "Molecular cloning and characterization of novel chicken survivin
RT splicing variants."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF377324; AA056309.1;
DR EMBL: AF377322; AA056307.1;
DR SEQUENCE 59 AA; 5716 MW; 79F59B7DA6C2A2CA CRC64;

Query Match 6.3%; Score 9; DB 13; Length 59;
Best Local Similarity 100.0%; Pred.No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 EGCACTPER 37
Db 31 EGCACTPER 39

RESULT 11
Q90WU8 PRELIMINARY; PRT; 128 AA.
ID Q90WU8;
AC Q90WU8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Survivin 2.
GN BIR5B.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Inohara N., Nunez G.;
RT MEDLINE=20373792; PubMed=10917738;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish."
RL Cell Death Differ. 7:509-510(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Inohara N., Nunez G.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY057058; AA118251.1;
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0008189; F:apoptosis inhibitor activity; IEA.
DR GO: GO:0006916; P:anti-apoptosis; IEA.
DR InterPro: IPR001370; BIR.
DR Pfam: PF00653; BIR; 1.
DR SMART: SM00238; BIR; 1.
DR PROSITE: PS00143; BIR_REPEAT_2; 1.
SQ SEQUENCE 128 AA; 15294 MW; AB496849082DDC66 CRC64;

Query Match 6.3%; Score 9; DB 13; Length 128;
Best Local Similarity 100.0%; Pred.No. 0.31;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 ELEGWEPDD 71
Db 52 ELEGWEPDD 60

```

```

Matches      8;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

Qy      135 IEQLAMD 142
Db      41 IEQLAMD 48

RESULT 14
Q82YH3      PRELIMINARY;      PRT;      197 AA.
ID      Q82YH3
AC      Q82YH3;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical protein.
GN      SAP1.7.
OS      Streptomyces avermitilis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Streptomycineae; Streptomycetaceae; Streptomyces.
OX      NCBI_TaxID=33903;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=MA-4880 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX      MEDLINE=22608306; PubMed=12692562;
RA      Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA      Sakaki Y., Hattori M., Omura S.;
RT      "Complete genome sequence and comparative analysis of the industrial
RT      microorganism Streptomyces avermitilis.";
RL      Nat. Biotechnol. 21:526-531(2003).
DR      GO; GO:005645; BAC75291.1; -.
DR      GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW      Hypothetical protein; Plasmid; Complete proteome.
SQ      SEQUENCE 197 AA; 21399 MW; 059D770B58CA2F75 CRC64;

Query Match      5.6%; Score 8; DB 16; Length 197;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches      8;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

Qy      2 GAPTLPFA 9
Db      113 GAPTLPFA 120

RESULT 15
Q9F8X4      PRELIMINARY;      PRT;      234 AA.
ID      Q9F8X4
AC      Q9F8X4;
DT      01-MAR-2001 (TrEMBLrel. 16, Created)
DT      01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      TrER.
GN      TrER.
OS      Pseudomonas fluorescens.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Pseudomonadaceae; Pseudomonas.
OX      NCBI_TaxID=294;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 17400;
RX      MEDLINE=98027372; PubMed=9361421;
RA      Gaballa A., Abeysinghe P.D., Ulrich G., Matthijs S., De Greve H.,
RA      Cornelis P., Koedam N.;
RT      "Trehalose induces antagonism towards Pythium debaryanum in
RT      Pseudomonas fluorescens ATCC 17400.";
RL      Appl. Environ. Microbiol. 63:4340-4345(1997).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 17400;
RX      MEDLINE=21030915; PubMed=11191810;
RA      Matthijs S.L.C., Koedam N., Cornelis P., De Greve H.;
RT      "The trehalose operon of Pseudomonas fluorescens ATCC 17400.";
RL      Res. Microbiol. 151:845-851(2000).

Matches      8;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

Qy      62 KELEGWEP 69
Db      61 KELEGWEP 68

RESULT 13
Q94EC5      PRELIMINARY;      PRT;      195 AA.
ID      Q94EC5
AC      Q94EC5;
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      P0002B05.14 protein.
GN      P0002B05.14.
OS      Oryza sativa (Rice).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Ehrhartoideae; Oryzeae; Oryza.
OX      NCBI_TaxID=4530;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=cv. Nipponbare;
RA      Sakaki T., Matsumoto T., Yamamoto K.;
RT      "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT      clone:P0002B05.";
RL      Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF003141; BAB63494.1; -.
DR      Gramene; Q94EC5; -.
SQ      SEQUENCE 195 AA; 21153 MW; CD2BFC4B9D005A1A CRC64;

Query Match      5.6%; Score 8; DB 10; Length 195;
Best Local Similarity 100.0%; Pred. No. 5;
Matches      8;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

Qy      62 KELEGWEP 69
Db      61 KELEGWEP 68

RESULT 12
Q90WU9      PRELIMINARY;      PRT;      142 AA.
ID      Q90WU9
AC      Q90WU9;
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Survivin 1.
GN      BIRSA.
OS      Brachydanio rerio (Zebrafish) (Danio rerio).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC      Cyprinidae; Danio.
OX      NCBI_TaxID=7955;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=2037392; PubMed=10917738;
RA      Inohara N., Nunez G.;
RT      "Genes with homology to mammalian apoptosis regulators identified in
RT      zebrafish.";
RL      Cell Death Differ. 7:509-510(2000).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Inohara N., Nunez G.;
RL      Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY057057; AAL18250.1; -.
DR      GO; GO:0005622; C:intracellular; IEA.
DR      GO; GO:0008189; F:apoptosis inhibitor activity; IEA.
DR      GO; GO:0008916; P:anti-apoptosis; IEA.
DR      InterPro; IPR001370; BIR.
DR      Pfam; PF00653; BIR; 1.
DR      SMART; SM00238; BIR; 1.
DR      PROSITE; PS0143; BIR_REPEAT_2; 1.
DR      PROSITE; PS0143; BIR_REPEAT_2; 1.
SQ      SEQUENCE 142 AA; 16412 MW; 71AED239452B5C78 CRC64;

Query Match      5.6%; Score 8; DB 13; Length 142;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches      8;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

Qy      62 KELEGWEP 69
Db      61 KELEGWEP 68

```

CC -!- SIMILARITY: BELONGS TO THE GNTR FAMILY OF TRANSCRIPTIONAL REGULATORS.

DR EMBL: AF229829; AAC31030.1; -.  
 DR GO: GO:0005622; C:intracellular; IEA.  
 DR GO: GO:0003700; F:transcription factor activity; IEA.  
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR GO: GO:0006350; P:transcription; IEA.  
 DR InterPro: IPR000524; HTH\_Gntr.  
 DR Pfam: PF00392; gntr; 1.  
 DR PRINTS: PR00035; HTHGNTR.  
 DR SMART: SM00345; HTH GNTR; 1.

KW DNA-binding; Transcription; Transcription regulation.  
 SQ SEQUENCE 234 AA; 26716 MW; 538FC11BFBD2C32B CRC64;

Query Match 5.6%; Score 8; DB 2; Length 234;  
 Best Local Similarity 100.0%; Pred. No. 5.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 VRRRAIEQL 138  
 |||||  
 Db 43 VRRRAIEQL 50

## RESULT 16

Q95YA6 PRELIMINARY; PRT; 491 AA.

AC Q95YA6;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein F41G4.5.  
 GN F41G4.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA Wilson R.;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).

[2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Miller N., Gattung S.;  
 RT "The sequence of C. elegans cosmid F41G4.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

[3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.H.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

[5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

[6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

[7]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;

RA Waterston R.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RN [8]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

RN [9]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilson R.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.

RN [10]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilson R.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.

RN [11]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilson R.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC006638; AAK85485.2; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 491 AA; 56030 MW; C1C08713B0A88486 CRC64;

Query Match 5.6%; Score 8; DB 5; Length 491;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GAPTLPFA 9  
 |||||  
 Db 392 GAPTLPFA 399

## RESULT 17

P72747 PRELIMINARY; PRT; 706 AA.

AC P72747;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein slr1103.  
 GN SLR1103.  
 OS Synechocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

OX NCBI\_TaxID=1148;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97061201; PubMed=9905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
 RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
 RA Tabata S.;

RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 DR EMBL: D90900; BAA16762.1; -.  
 DR PIR: S74610; S74610.

DR InterPro: IPR001633; EAL.  
 DR InterPro: IPR00253; FFA.  
 DR InterPro: IPR00160; GGDEF.  
 DR InterPro: IPR008984; SMAD\_FHA.  
 DR Pfam: PF00563; EAL; 1.  
 DR Pfam: PF00498; FHA; 1.

DR Pfam: PF00590; GGDEF; 1.  
 DR SMART: SM00267; DUF1; 1.  
 DR SMART: SM00052; DUF2; 1.  
 DR SMART: SM00240; FHA; 1.  
 DR TIGRFAMs: TIGR00254; GGDEF; 1.

DR PROSITE: PS50883; EAL; 1.  
 DR PROSITE: PS50006; FHA DOMAIN; 1.  
 DR PROSITE: PS50887; GGDEF; 1.  
 KW Hypothetical protein; Complete proteome.

RN Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE 706 AA; 79830 MW; 5B909B4AA5DAC300 CRC64;



Query Match 5.6%; Score 8; DB 16; Length 706;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 132 REAIEOLA 139

Db 585 RRAIEOLA 592

RESULT 18

O18265 PRELIMINARY; PRT; 875 AA.  
AC O18265;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE ZC15.1 protein.  
GN ZC15.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Feloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Basham V.M.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL; Z93396; CA807709.1; -.  
DR FIR; T27495; T27495.  
DR WormPep; ZC15.1; CE16683.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR InterPro; IPR001878; Znf CCHC.  
DR SMART; SM00343; Znf C2HC; 1.  
SQ SEQUENCE 875 AA; 100139 MW; 9B12345834B6AB84 CRC64;

Query Match 5.6%; Score 8; DB 5; Length 875;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GAPTLPAA 9

Db 433 GAPTLPAA 440

RESULT 19

O8ILL7 PRELIMINARY; PRT; 1621 AA.  
AC O8ILL7;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Hypothetical protein.  
GN PF14\_0226.  
OS Plasmodium falciparum (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=36329;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=3D7;  
RX MEDLINE=22255705; PubMed=12368864;  
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,  
RA Carlton J.N., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,  
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,  
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,  
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,  
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,  
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,

RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,  
RA Fraser C.M., Barrell B.;  
RT "Genome sequence of the human malaria parasite Plasmodium  
falciparum";  
RL Nature 419:498-511(2002).  
DR EMBL; AE014819; AAN36839.1; -.  
DR InterPro; IPR001646; Speptide\_repeat.  
DR Pfam; PF08085; Pentapeptide; 2.  
KW Hypothetical protein.  
SQ SEQUENCE 1621 AA; 192605 MW; B4492093FDAD3A4E CRC64;

Query Match 5.6%; Score 8; DB 5; Length 1621;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 EEHKHSS 82

Db 1019 EEHKHSS 1026

RESULT 20

Q07290 PRELIMINARY; PRT; 1822 AA.  
AC Q07290;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Epi\* protein.  
GN EPI\*.  
OS Streptococcus suis.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1307;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1890;  
RX MEDLINE=93328288; PubMed=8335363;  
RA Smith H.E., Reek F.H., Vecht U., Gielkens A.L.J., Smits M.A.;  
RT "Repeats in an extracellular protein of weakly pathogenic strains of  
Streptococcus suis type 2 are absent in pathogenic strains."  
RL Infect. Immun. 61:3318-3326(1993).  
DR EMBL; X71880; CAA50714.1; -.  
DR PIR; S33441; S33441.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR005877; Gpos YSIK.  
DR Pfam; PF04650; YSIK signal; 1.  
DR TIGRfams; TIGR01168; YSIK signal; 1.  
SQ SEQUENCE 1822 AA; 192632 MW; 3838960C77641D7D CRC64;

Query Match 5.6%; Score 8; DB 2; Length 1822;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 AKNKIAKE 116

Db 1057 AKNKIAKE 1064

RESULT 21

O57137 PRELIMINARY; PRT; 66 AA.  
AC O57137;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Paracoccus denitrificans.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;  
OC Rhodobacteraceae; Paracoccus.  
OX NCBI\_TaxID=266;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95370157; PubMed=7642505;

```

RA Van Spanning R.J., Reijnders W.N., Stouthamer A.H.;
RT "Integration of heterologous DNA into the genome of Paracoccus
RT denitrificans is mediated by a family of IS1248-related elements and a
RT second type of integrative recombination event.";
RL J. Bacteriol. 177:4772-4778 (1995).
DR ENBL; U08856; AAA82147.1; -
DR ENBL; U08864; AAC43508.1; -
KW Hypothetical protein.
SQ SEQUENCE 66 AA; 7025 MW; 1865DD2D0F8316EC3 CRC64;

Query Match 4.9%; Score 7; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 APTLPPA 9
Db 33 APTLPPA 39

RESULT 22
Q8DS63 PRELIMINARY; PRT; 102 AA.
AC Q8DS63;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Conserved hypothetical protein.
GN SMO.1972C.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=2295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
DR ENBL; AE015020; AAN59577.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 102 AA; 11678 MW; E4B1AA4265E0F554 CRC64;

Query Match 4.9%; Score 7; DB 16; Length 102;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 VKQFEE 95
Db 27 VKQFEE 33

RESULT 23
Q8PKK9 PRELIMINARY; PRT; 118 AA.
AC Q8PKK9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein XAC2162.
GN XAC2162.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,

RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fortmighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.I.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463 (2002).
DR ENBL; AE011853; AAM37015.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 118 AA; 12295 MW; DCEC804DDB651B0 CRC64;

Query Match 4.9%; Score 7; DB 16; Length 118;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 APTLPPA 9
Db 44 APTLPPA 50

RESULT 24
Q9ICW7 PRELIMINARY; PRT; 125 AA.
AC Q9ICW7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA replicase (fragment).
OS Kalanchoe latent virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.
OX NCBI_TaxID=132477;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PV-0290;
RX MEDLINE=21342909; PubMed=11450944;
RA Nicolaissen M., Nielsen S.L.;
RT "Analysis of the triple gene block and coat protein sequences of two
RT strains of Kalanchoe Latent Carlavirus.";
RL Virus Genes 22:265-270 (2001).
DR ENBL; AJ293571; CAB97502.1; -
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR001788; RNA_dep_RNAPol2.
DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
FT NON_TER
SQ SEQUENCE 125 AA; 14327 MW; 20867AB30B7BA945 CRC64;

Query Match 4.9%; Score 7; DB 12; Length 125;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 IAKETNN 119
Db 57 IAKETNN 63

RESULT 25
Q18231 PRELIMINARY; PRT; 130 AA.
AC Q18231;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```

```

RL Virus Genes 22:265-270(2001).
DR EMEL; AJ293570; CAB97496.1; -.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR001788; RNA_dep_RNAPol2.
DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
FT NON TER 1
SQ SEQUENCE 130 AA; 14842 MW; 770B214D0D6CD7A0 CRC64;

Query Match 4.9%; Score 7; DB 12; Length 130;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 IAKETNN 119
Db 62 IAKETNN 68
|||||

RESULT 27
Q8QXX8 PRELIMINARY; PRT; 149 AA.
ID Q8QXX8
AC Q8QXX8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA-dependent RNA-polymerase (Fragment).
OS Garlic latent virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.
OX NCBI_TaxID=12458;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VH3;
EX MEDLINE=21578618; PubMed=11722009;
RT Chen J.; Chen J.P.; Adams M.J.;
RT "Molecular characterisation of a complex mixture of viruses in garlic
RT with mosaic symptoms in China.";
RT Arch. Virol. 146:1841-1853(2001).
DR EMEL; AJ292228; CAC83692.1; -.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR001788; RNA_dep_RNAPol2.
DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
FT NON TER 1
SQ SEQUENCE 149 AA; 17307 MW; BE6B5747657A7F54 CRC64;

Query Match 4.9%; Score 7; DB 12; Length 149;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 IAKETNN 119
Db 81 IAKETNN 87
|||||

RESULT 28
Q8ZJX0 PRELIMINARY; PRT; 152 AA.
ID Q8ZJX0
AC Q8ZJX0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative cytoplasmic protein.
GN STM4549.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSCi412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;

```

```

RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Ngwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
DR EMBL; AE008914; AAL23367.1; -.
DR InterPro; IPR007214; YbaK.
DR Pfam; PF04073; YbaK; 1.
DR Hypothetical protein; Complete proteome.
KW Hypothetical protein; Complete proteome.
SQ
SEQUENCE 152 AA; 16273 MW; 1DE1B50810C8653 CRC64;

Query Match 4.9%; Score 7; DB 16; Length 152;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 TPERMAE 40
Db 135 TPERMAE 141
|||||
RESULT 29
Q6ZOV6 PRELIMINARY; PRT; 152 AA.
ID Q8ZOV6
AC Q8ZOV6;
RT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DR 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein STY4902.
DN STY4902 OR T4592.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RA "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RA "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 195:2330-2337(2003).
DR EMBL; AL627284; CAD03387.1; -.
DR EMBL; AE016849; AAO70207.1; -.
DR InterPro; IPR007214; YbaK.
DR Pfam; PF04073; YbaK; 1.
KW Hypothetical protein; Complete proteome.
SQ
SEQUENCE 152 AA; 16229 MW; 1DE1B3B833EB3202 CRC64;

Query Match 4.9%; Score 7; DB 16; Length 152;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 TPERMAE 40
Db 135 TPERMAE 141
|||||

```

```

Db 135 TPERMAE 141
RESULT 30
Q89D81 PRELIMINARY; PRT; 155 AA.
ID Q89D81
AC Q89D81;
RT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DR 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BLR7561 protein.
DN BLR7561.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashina K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005982; BAC52826.1; -.
DR InterPro; IPR007214; YbaK.
DR Pfam; PF04073; YbaK; 1.
KW Complete proteome.
SQ
SEQUENCE 155 AA; 16422 MW; 469A81F631A1A17F CRC64;

Query Match 4.9%; Score 7; DB 16; Length 155;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 TPERMAE 40
Db 135 TPERMAE 141
|||||
RESULT 31
Q9QBX0 PRELIMINARY; PRT; 171 AA.
ID Q9QBX0
AC Q9QBX0;
RT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DR 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Replicase (Fragment).
GN REP.
OS Lily latent virus (LiLV).
OX Viruses; ssRNA positive-strand viruses, no DNA stage; Carlaviruses.
OX NCBI_TaxID=92693;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Lilv-KK;
RA Ryu J.H., Park H.W., Park W.M., Lee S.Y., Ryu K.H.;
RT "Molecular analysis of the 3'-terminal region of lily latent
RT Carlavirus from Lilium lancifolium.";
RL J. Plant Pathol. 16:231-235(2000).
DR EMBL; AJ131812; CAB57958.1; -.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:Viral genome replication; IEA.
DR InterPro; IPR001788; RNA_dep_RNAPol2.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
FT NON_TER
SQ
SEQUENCE 171 AA; 19849 MW; 246D7C9729465F8 CRC64;

Query Match 4.9%; Score 7; DB 12; Length 171;

```

Best Local Similarity 100.0%; Pred. No. 52;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 TAKETNN 119  
Db 103 TAKETNN 109

RESULT 32  
Q84F94 PRELIMINARY; PRT; 175 AA.  
AC Q84F94;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE HslV-like protein.  
OS Myxococcus xanthus.  
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;  
OC Cytochrome P450; Myxococcaceae; Myxococcus.  
OX NCBI\_TaxID=34;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hartzell P.L., Youderian P.A.;  
RT "Identification of genes required for adventurous gliding motility in  
RT Myxococcus xanthus with the transposable element mariner.";  
RL Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AY204472; AAC22924.1; -.  
DR GO; GO:0005839; C:proteasome core complex (sensu Eukarya); IEA.  
DR GO; GO:0004175; F:endoropeptidase activity; IEA.  
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.  
DR InterPro; IPR001353; Peptidase\_T1.  
DR Pfam; PF00227; Proteasome; 1.  
SQ SEQUENCE 175 AA; 19060 MW; D72BFCB837A6950A CRC64;

Query Match 4.9%; Score 7; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 TAKKVR 133  
Db 34 TAKKVR 40

RESULT 33  
Q8XKR3 PRELIMINARY; PRT; 178 AA.  
AC Q8XKR3;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Rubrerythrin.  
GN RUBY OR CPE1331.  
OS Clostridium perfringens.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1502;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=13 / Type A;  
RX MEDLINE=21664373; PubMed=11792842;  
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,  
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;  
RT "Complete genome sequence of Clostridium perfringens, an anaerobic  
RT flesh-eater.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).  
DR EMBL; AF003190; BAB81037.1; -.  
DR GO; GO:0046872; F:metal ion binding; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR009040; Ferritin like.  
DR InterPro; IPR004039; Rubredox.  
DR InterPro; IPR003251; Rubrerythrin.  
DR Pfam; PF02915; Rubrerythrin; 1.  
DR ProDom; PD006196; Rubrerythrin; 1.

DR PROSITE; PS50905; FERRITIN LIKE; 1.  
DR PROSITE; PS50903; RUBREDOXIN LIKE; 1.  
KW Complete proteome.  
SQ SEQUENCE 178 AA; 21033 MW; F55A85ADB0F56A01 CRC64;  
Query Match 4.9%; Score 7; DB 16; Length 178;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 KEFEETA 128  
Db 91 KEFEETA 97

RESULT 34  
Q834K3 PRELIMINARY; PRT; 182 AA.  
AC Q834K3;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Heat shock protein HslV.  
GN HSLV OR EF1647.  
OS Enterococcus faecalis (Streptococcus faecalis).  
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
OX NCBI\_TaxID=1351;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=V583 / ATCC 700802;  
RX MEDLINE=22550857; PubMed=12663927;  
RA Paulsen I.T., Banerjee L., Myers G.S.A., Gill S.R., Heidelberg J.F.,  
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Beanan M.,  
RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,  
RA Doughterty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,  
RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,  
RA Urtuback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;  
RT "Role of mobile DNA in the evolution of vancomycin-resistant  
RT Enterococcus faecalis.";  
RL Science 299:2071-2074(2003).  
DR EMBL; AB016952; AAC1425.1; -.  
DR TIGR; EF1647; -.  
DR GO; GO:0005839; C:proteasome core complex (sensu Eukarya); IEA.  
DR GO; GO:0004175; F:endoropeptidase activity; IEA.  
DR GO; GO:0003773; F:heat shock protein activity; IEA.  
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.  
DR InterPro; IPR001353; Peptidase\_T1.  
DR Pfam; PF00227; Proteasome; 1.  
DR Heat shock; Complete proteome.  
KW Heat shock; Complete proteome.  
SQ SEQUENCE 182 AA; 19712 MW; 76C1D6ABA7DE2237 CRC64;

Query Match 4.9%; Score 7; DB 16; Length 182;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 TAKKVR 133  
Db 39 TAKKVR 45

RESULT 35  
Q9JMX9 PRELIMINARY; PRT; 199 AA.  
AC Q9JMX9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE Cag-Z.  
GN CAG-Z.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Helicobacteraceae; Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=CCUG 17874;
RX MEDLINE=97121442; PubMed=8962108;
RA Censini S., Lange C., Xiang Z., Crabtree J.E., Ghiara P.,
RA Borodovsky M., Rappuoli R., Covacci A.;
RT "cag, a pathogenicity island of Helicobacter pylori, encodes type I-
RT specific and disease-associated virulence factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:14648-14653 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CCUG 17874;
RX MEDLINE=20150112; PubMed=10684850;
RA Covacci A., Rappuoli R.;
RT "Tyrosine-phosphorylated bacterial proteins: Trojan horses for the
RT host cell.";
RL J. Exp. Med. 191:587-592 (2000).
DR EMBL; AF282852; AAF80197.1; -.
SQ SEQUENCE 199 AA; 23044 MW; 8D1F2D5AB4B6650B CRC64;

Query Match 4.9%; Score 7; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 ETLTGEF 101
Db 98 ETLTGEF 104

RESULT 36
Q025261 PRELIMINARY; PRT; 199 AA.
AC O25261;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CAG pathogenicity island protein (CAG6).
GN HP0526
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=5252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547 (1997).
DR EMBL; AE000566; AAD07593.1; -.
DR FIR; F64585; F64585.
DR TIGR; HP0526; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 199 AA; 23017 MW; 68F2F32BAF61C814 CRC64;

Query Match 4.9%; Score 7; DB 16; Length 199;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 ETLTGEF 101
Db 98 ETLTGEF 104

RESULT 37

```

```

Q092LV1 PRELIMINARY; PRT; 199 AA.
AC Q092LV1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CAG island protein.
GN JHP0475.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180 (1999).
DR EMBL; AE001481; AAD06046.1; -.
DR FIR; H71927; H71927.
KW Complete proteome.
SQ SEQUENCE 199 AA; 23126 MW; 30C7306F38E450BF CRC64;

Query Match 4.9%; Score 7; DB 16; Length 199;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 ETLTGEF 101
Db 98 ETLTGEF 104

RESULT 38
Q08D904 PRELIMINARY; PRT; 219 AA.
AC Q08D904;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN V12809.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.-S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016806; AAO11147.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 219 AA; 24982 MW; 70D87332AA4780406 CRC64;

Query Match 4.9%; Score 7; DB 16; Length 219;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 132 RRAIEQL 138
Db 17 RRAIEQL 23

RESULT 39
Q07U7M6 PRELIMINARY; PRT; 224 AA.
ID Q07U7M6

```

```

AC Q0UTM6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical.
GN SYN0955.
OS Synechococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A. PubMed=12917641;
RX MEDLINE=22825697;
RA Palenik B., Brahamsha B., Larimer F.W., Land M., Hauser L., Chain P.,
RA Lamerdin J., Regala W., Allen E.E., McCarron J., Paulsen I.,
RA Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synechococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; EX569691; CAB07470.1; -.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 224 224 Potential.
SQ SEQUENCE 224 AA; 25402 MW; D22937524F061D2B CRC64;

Query Match 4.9%; Score 7; DB 16; Length 224;
Best Local Similarity 100.0%; Pred. No. 65; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

Oy 104 LDRERAK 110
Db 92 LDRERAK 98
|||||

RESULT 40
O58403 PRELIMINARY; PRT; 228 AA.
AC O58403;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 228AA long hypothetical aspartate racemase.
GN PH0670.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98344137; PubMed=9679194;
RA Kwarabavasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL; AF000003; BAA29761.1; -.
DR FIR; G71112; G7112.
DR PDB; 1JFL; 25-DEC-02.
DR GO; GO:0016855; F:racemase and epimerase activity, acting on . . .; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0001920; Asp/Glu_rac.
DR InterPro; IPR001920; Asp_rac.
DR InterPro; IPR004380; Asp_rac.
DR Pfam; PF01177; Asp_Glu_rac; 1.
DR TIGRFAMS; TIGR00035; asp_rac; 1.
DR PROSITE; PS00923; ASP_GLU_RACEMASE_1; 1.
DR PROSITE; PS00924; ASP_GLU_RACEMASE_2; 1.
KW Complete proteome.
SQ SEQUENCE 228 AA; 25157 MW; DDA2CBEB39682CF CRC64;

Query Match 4.9%; Score 7; DB 17; Length 228;
Best Local Similarity 100.0%; Pred. No. 66; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

Oy 125 BETAKKV 131
Db 105 BETAKKV 111
|||||

RESULT 42
Q37965 PRELIMINARY; PRT; 248 AA.
ID Q37965
AC Q37965;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Antiterminator.
GN Q OR GENE 40.
OS Bacteriophage N15.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=40631;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96213032; PubMed=8631680;
RA Lobočka M.B., Svarchevsky A.N., Rybchin V.N., Yarmolinsky M.B.;
RT "Characterization of the primary immunity region of the Escherichia
RT coli linear plasmid prophage N15.";
RL J. Bacteriol. 178:2902-2910(1996).
RN [2]

```

```

Best Local Similarity 100.0%; Pred. No. 66; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

Oy 125 BETAKKV 131
Db 105 BETAKKV 111
|||||

RESULT 41
Q9UYV1 PRELIMINARY; PRT; 228 AA.
ID Q9UYV1
AC Q9UYV1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Aspartate racemase.
GN RACD-1 OR PYRAB13760 OR PAB0912.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GS / Orsay;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RT structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ248287; CAB50281.1; -.
DR PIR; D75048; D75048.
DR GO; GO:0016855; F:racemase and epimerase activity, acting on . . .; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001920; Asp/Glu_rac.
DR InterPro; IPR004380; Asp_rac.
DR Pfam; PF01177; Asp_Glu_rac; 1.
DR TIGRFAMS; TIGR00035; asp_rac; 1.
DR PROSITE; PS00923; ASP_GLU_RACEMASE_1; 1.
DR PROSITE; PS00924; ASP_GLU_RACEMASE_2; 1.
KW Complete proteome.
SQ SEQUENCE 228 AA; 25097 MW; 88B261B57A1D18C6 CRC64;

Query Match 4.9%; Score 7; DB 17; Length 228;
Best Local Similarity 100.0%; Pred. No. 66; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

Oy 125 BETAKKV 131
Db 105 BETAKKV 111
|||||

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=N15CML;
RA Loboeka M.B.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Hendrix R.W., Ravin V.K., Casjens S.R., Ford M.E., Ravin N.V.,
RA Smirnov I.K.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; U45242; AB881658.1; -.
DR EMBL; AF064539; AAC19067.1; -.
DR PIR; T13126; T13126.
DR GO; GO:0005743; C:Mitochondrial inner membrane; IEA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001993; Mitoch carrier.
DR PROSITE; PS00215; MITOCH CARRIER; 1.
SQ SEQUENCE 246 AA; 28817 MW; 3CBEA8CC0388C7D3 CRC64;

Query Match 4.9%; Score 7; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 RRAIEQL 138
Db 91 RRAIEQL 97

RESULT 43
Q822K6 PRELIMINARY; PRT; 250 AA.
ID Q822K6
AC Q822K6
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Phosphoglycerate-binding transcriptional regulator, RpiR family,
DE putative.
GN EF3049.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA Dougherty S., DeBoy R.T., Durkin A.S., Kolonay J., Madupu R.,
RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA Uterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis";
RL Science 299:2071-2074 (2003).
DR EMBL; AE016956; AA082731.1; -.
DR TIGR; BF3049; -.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001347; SIS.
DR Pfam; PF01380; SIS; 1.
DR Complete proteome.
SQ SEQUENCE 250 AA; 28455 MW; 7C37D6590E2BC9AC CRC64;

Query Match 4.9%; Score 7; DB 16; Length 250;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 EETAKKV 131
Db 158 EETAKKV 164

RESULT 44

```

```

Q92TW2 PRELIMINARY; PRT; 253 AA.
ID Q92TW2
AC Q92TW2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative proteocatechuate 3,4-dioxygenase beta chain protein (EC
DE 1.13.11.3).
GN PCAH OR R81489 OR SM320577.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymB (megaplasmid 2).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21395508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puhler A.;
RT "The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894 (2001).
DR EMBL; AL603647; CAC49889.1; -.
DR PIR; A96028; A96028.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0016702; F:oxidoreductase activity, acting on single d. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR InterPro; IPR000627; Dioxxygenase.
DR Pfam; PF00775; Dioxxygenase; 1.
DR PROSITE; PS00083; INTRADIOL DIOXYGENAS; 1.
DE Dioxxygenase; Oxidoreductase; Plasmid; Hypothetical protein;
DE Complete proteome.
SQ SEQUENCE 253 AA; 28522 MW; 8D3A91173A144606 CRC64;

Query Match 4.9%; Score 7; DB 16; Length 253;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 RRAIEQL 138
Db 209 RRAIEQL 215

RESULT 45
Q8EBM3 PRELIMINARY; PRT; 255 AA.
ID Q8EBM3
AC Q8EBM3
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transcriptional regulator, AraC/XylS family.
GN SO3488.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Dougherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Imprim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Uterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealeon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium

```



RT Shewanella oneidensis.",  
RL Nat. Biotechnol. 20:1118-1123(2002).  
DR EMBL; AE015785; AAN56481.1; -.  
DR TIGR; SO3488; -.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR DR GO; GO:0003700; P:transcription factor activity; IEA.  
DR GO; GO:0006355; R:regulation of transcription, DNA-dependent; IEA.  
DR IncerPro; IPR003313; ASAC binding.  
DR InterPro; IPR000005; HTHARAC.  
DR Pfam; PF02311; Arac binding; 1.  
DR Pfam; PF00165; HTH ARAC; 2.  
DR PRINTS; PR00032; HTHARAC.  
DR SMART; SMO0342; HTH ARAC; 1.  
DR PROSITE; PS00041; HTH ARAC FAMILY\_1; 1.  
DR PROSITE; PS01124; HTH\_ARAC\_FAMILY\_2; 1.  
DR Complete proteome.  
SQ SEQUENCE 255 AA; 28797 MW; B2B48B542E4235E CRC64;  
  
Query Match 4.9%; Score 7; DB 16; Length 255;  
Best Local Similarity 100.0%; Pred.No. 72;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 134 AIEQLAA 140  
|||  
DB 213 AIEQLAA 219  
|||  
  
RESULT 46  
Q9M521 PRELIMINARY; PRT; 265 AA.  
ID Q9M521  
AC Q9M521;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DI 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DE Pinorexinol-laricresinol reductase TH1.  
OS Tsuga heterophylla (Western hemlock).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Tsuga.  
OX NCBI\_TaxID=3359;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Stem;  
RX MEDLINE=99167523; PubMed=10066819;  
RA Gang D.R., Kasahara H., Xia Z.Q., Vander Minnsbrugge K., Bauw G., Boerjan W., Van Montagu M., Davin L.B., Lewis N.G.; "Evolution of plant defense mechanisms. Relationships of phenylcoumaran benzylic ether reductases to pinorexinol-laricresinol and isoflavone reductases."; J. Biol. Chem. 274:7516-7527(1999).  
[2]  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Stem;  
RX MEDLINE=99167523; PubMed=10066819;  
RA Gang D.R., Kasahara H., Xia Z.Q., Davin L.B., Lewis N.G.; Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF242501; AAF64184.1; -.  
GO; GO:0003824; F:catalytic activity; IEA.  
DR InterPro; IPR003866; Isoflav\_reduc.  
DR Pfam; PF02716; Isoflavone\_redu; 1.  
SQ SEQUENCE 265 AA; 29987 MW; 7BCFB9EA09BA3CE CRC64;  
  
Query Match 4.9%; Score 7; DB 10; Length 265;  
Best Local Similarity 100.0%; Pred.No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 130 KVRRAIE 136  
|||  
DB 93 KVRRAIE 99  
|||  
  
RESULT 47  
C28569 PRELIMINARY; PRT; 271 AA.  
ID C28569  
AC C28569;  
RT Shewanella oneidensis.",  
RL Nat. Biotechnol. 20:1118-1123(2002).  
DR EMBL; AE015785; AAN56481.1; -.  
DR TIGR; SO3488; -.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR DR GO; GO:0003700; P:transcription factor activity; IEA.  
DR GO; GO:0006355; R:regulation of transcription, DNA-dependent; IEA.  
DR IncerPro; IPR003313; ASAC binding.  
DR InterPro; IPR000005; HTHARAC.  
DR Pfam; PF02311; Arac binding; 1.  
DR Pfam; PF00165; HTH ARAC; 2.  
DR PRINTS; PR00032; HTHARAC.  
DR SMART; SMO0342; HTH ARAC; 1.  
DR PROSITE; PS00041; HTH ARAC FAMILY\_1; 1.  
DR PROSITE; PS01124; HTH\_ARAC\_FAMILY\_2; 1.  
DR Complete proteome.  
SQ SEQUENCE 255 AA; 28797 MW; B2B48B542E4235E CRC64;  
  
Query Match 4.9%; Score 7; DB 16; Length 255;  
Best Local Similarity 100.0%; Pred.No. 72;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 134 AIEQLAA 140  
|||  
DB 213 AIEQLAA 219  
|||  
  
RESULT 46  
Q9M521 PRELIMINARY; PRT; 265 AA.  
ID Q9M521  
AC Q9M521;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DI 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DE Pinorexinol-laricresinol reductase TH1.  
OS Tsuga heterophylla (Western hemlock).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Tsuga.  
OX NCBI\_TaxID=3359;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Stem;  
RX MEDLINE=99167523; PubMed=10066819;  
RA Gang D.R., Kasahara H., Xia Z.Q., Vander Minnsbrugge K., Bauw G., Boerjan W., Van Montagu M., Davin L.B., Lewis N.G.; "Evolution of plant defense mechanisms. Relationships of phenylcoumaran benzylic ether reductases to pinorexinol-laricresinol and isoflavone reductases."; J. Biol. Chem. 274:7516-7527(1999).  
[2]  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Stem;  
RX MEDLINE=99167523; PubMed=10066819;  
RA Gang D.R., Kasahara H., Xia Z.Q., Davin L.B., Lewis N.G.; Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF242501; AAF64184.1; -.  
GO; GO:0003824; F:catalytic activity; IEA.  
DR InterPro; IPR003866; Isoflav\_reduc.  
DR Pfam; PF02716; Isoflavone\_redu; 1.  
SQ SEQUENCE 265 AA; 29987 MW; 7BCFB9EA09BA3CE CRC64;  
  
Query Match 4.9%; Score 7; DB 10; Length 265;  
Best Local Similarity 100.0%; Pred.No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 130 KVRRAIE 136  
|||  
DB 93 KVRRAIE 99  
|||  
  
RESULT 47  
C28569 PRELIMINARY; PRT; 271 AA.  
ID C28569  
AC C28569;

```

DR InterPro; IPR003866; Isoflav_reduct.
DR Pfam; PF02716; Isoflavone_redu; 1.
SQ SEQUENCE 308 AA; 33464 MW; F3CF7A3D123D0251 CRC64;

Query Match
Best Local Similarity 4.9%; Score 7; DB 10; Length 308;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAIE 136
DB 136 KVRRAIE 142
|||||

RESULT 49
Q9M522 PRELIMINARY; PRT; 308 AA.
ID Q9M522;
AC Q9M522;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phenylcoumaran benzylic ether reductase homolog TH7.
OS Tsuga heterophylla (Western hemlock).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Tsuga.
NCBI_TaxID=3359;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stem;
RX MEDLINE=99167523; PubMed=10066819;
RA Gang D.R., Kasahara H., Xia Z.Q., Davin L.B., Lewis N.G.;
RA Boerjan W., Van Montagu M., Davin L.B., Lewis N.G.;
RT "Evolution of plant defense mechanisms. Relationships of
phenylcoumaran benzylic ether reductases to pinosresinol-lariciresinol
and isoflavone reductases."
RT J. Biol. Chem. 274:7516-7527(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Stem;
RA Gang D.R., Kasahara H., Xia Z.Q., Davin L.B., Lewis N.G.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF242439; AAF64179.1;
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR003866; Isoflav_reduct.
DR Pfam; PF02716; Isoflavone_redu; 1.
SQ SEQUENCE 308 AA; 33849 MW; DIA3FE25215A26DD CRC64;

Query Match
Best Local Similarity 4.9%; Score 7; DB 10; Length 308;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAIE 136
DB 136 KVRRAIE 142
|||||

RESULT 50
Q9M525 PRELIMINARY; PRT; 308 AA.
ID Q9M525;
AC Q9M525;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phenylcoumaran benzylic ether reductase homolog TH4.
OS Tsuga heterophylla (Western hemlock).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Tsuga.
NCBI_TaxID=3359;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stem;
RX MEDLINE=99167523; PubMed=10066819;
RA Gang D.R., Kasahara H., Xia Z.Q., Davin L.B., Lewis N.G.;
RA Boerjan W., Van Montagu M., Davin L.B., Lewis N.G.;
RT "Evolution of plant defense mechanisms. Relationships of
phenylcoumaran benzylic ether reductases to pinosresinol-lariciresinol
and isoflavone reductases."
RT J. Biol. Chem. 274:7516-7527(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Stem;
RA Gang D.R., Kasahara H., Xia Z.Q., Davin L.B., Lewis N.G.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF242439; AAF64182.1;
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR003866; Isoflav_reduct.
DR Pfam; PF02716; Isoflavone_redu; 1.
SQ SEQUENCE 308 AA; 33849 MW; DIA3FE25215A26DD CRC64;

Query Match
Best Local Similarity 4.9%; Score 7; DB 10; Length 308;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAIE 136
DB 136 KVRRAIE 142
|||||

RESULT 51
O81651 PRELIMINARY; PRT; 308 AA.
ID O81651;
AC O81651;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phenylcoumaran benzylic ether reductase.
GN DDCBER1.
OS Pinus taeda (loblolly pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
NCBI_TaxID=3352;
RN [1]
RP SEQUENCE FROM N.A.
RA Gang D.R., Kasahara H., Davin L.B., Lewis N.G.;
RT "Cloning of phenylcoumaran benzylic ether reductase from Pinus taeda
and heterologous expression of its recombinant protein in E. coli."
RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF081678; AAC32591.1;
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR003866; Isoflav_reduct.
DR Pfam; PF02716; Isoflavone_redu; 1.
SQ SEQUENCE 308 AA; 33565 MW; 07F9DD1F2CC8B134 CRC64;

Query Match
Best Local Similarity 4.9%; Score 7; DB 10; Length 308;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAIE 136
DB 136 KVRRAIE 142
|||||

RESULT 52
O3LL41 PRELIMINARY; PRT; 308 AA.
ID O3LL41;
AC O3LL41;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phenylcoumaran benzylic ether reductase Pfl.
OS Pinus taeda (loblolly pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
NCBI_TaxID=3352;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99167523; PubMed=10066819;

```

RA Gang D.R., Kasahara H., Xia Z.Q., Vander Mijnsbrugge K., Bauw G.,  
 RA Boerjan W., Van Montagu M., Davin L.B., Lewis N.G.;  
 RT "Evolution of plant defense mechanisms. Relationships of  
 RT phenylcoumaran benzylic ether reductases to pinoreinsol-lariciresinol  
 RT and isoflavone reductases.";  
 RL J. Biol. Chem. 274:7516-7527 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gang D.R., Kasahara H., Xia Z.Q., Davin L.B., Lewis N.G.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF242490; AAF64173.2; -;  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR InterPro; IPR003866; Isoflav\_reduct.  
 DR Pfam; PF02716; Isoflavone\_redu; 1.  
 SQ SEQUENCE 308 AA; 33524 MW; E93FA4456C7854SF CRC64;

Query Match 4.9%; Score 7; DB 10; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 84;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 KVRRAIE 136  
 |||||  
 Db 136 KVRRAIE 142

RESULT 53

Q9MS20 PRELIMINARY; PRT; 309 AA.

AC Q9MS20;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Pinoreinsol-lariciresinol reductase TH2.  
 OS Tsuga heterophylla (Western hemlock).  
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 CC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Tsuga.  
 OX NCBI\_TaxID=3359;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Stem;  
 RX MEDLINE=99167523; PubMed=10066819;  
 RA Gang D.R., Kasahara H., Xia Z.Q., Vander Mijnsbrugge K., Bauw G.,  
 RA Boerjan W., Van Montagu M., Davin L.B., Lewis N.G.;  
 RT "Evolution of plant defense mechanisms. Relationships of  
 RT phenylcoumaran benzylic ether reductases to pinoreinsol-lariciresinol  
 RT and isoflavone reductases.";  
 RL J. Biol. Chem. 274:7516-7527 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Stem;  
 RA Gang D.R., Kasahara H., Xia Z.Q., Davin L.B., Lewis N.G.;  
 RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF242502; AAF64185.1; -;  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR InterPro; IPR003866; Isoflav\_reduct.  
 DR Pfam; PF02716; Isoflavone\_redu; 1.  
 SQ SEQUENCE 309 AA; 34814 MW; 5ADBE69A362E3F3E1 CRC64;

Query Match 4.9%; Score 7; DB 10; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 KVRRAIE 136  
 |||||  
 Db 137 KVRRAIE 143

RESULT 54

Q9SD20 PRELIMINARY; PRT; 310 AA.

AC Q9SD20;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Isoflavone reductase homolog 2.  
 GN IPR2;  
 OS Glycine max (Soybean).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 OX NCBI\_TaxID=3847;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seed coat;  
 RA Tran M.H., Kufliu K., Gijzen M.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF202184; AAF1578.1; -;  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR InterPro; IPR003866; Isoflav\_reduct.  
 DR Pfam; PF02716; Isoflavone\_redu; 1.  
 SQ SEQUENCE 310 AA; 33939 MW; 2E76DFPA1BA199B0 CRC64;

Query Match 4.9%; Score 7; DB 10; Length 310;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 KVRRAIE 136  
 |||||  
 Db 137 KVRRAIE 143

RESULT 55

Q67551 PRELIMINARY; PRT; 312 AA.

AC Q67551;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE RNA replicase (Fragment).  
 OS Garlic latent virus.  
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlaviruses.  
 OX NCBI\_TaxID=12458;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tsuneyoshi T., Sumi S.;  
 RT "Differentiation among garlic viruses in mixed infections based on RT-  
 RT PCR procedures and direct tissue blotting immunoassays.";  
 RL Phytopathology 86:253-259 (1996).  
 DR EMBL; D11161; EAA01931.1; -;  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003688; F:RNA-directed RNA polymerase activity; IEA.  
 DR GO; GO:0006350; P:transcription; IEA.  
 DR GO; GO:0019079; P:viral genome replication; IEA.  
 DR InterPro; IPR001788; RNA\_dep\_RNAPol2.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR007094; RNA\_pol\_PSVir.  
 DR Pfam; PF00978; RNA\_dep\_RNAPol2; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 312 AA; 36230 MW; C79E6563DA488175 CRC64;

Query Match 4.9%; Score 7; DB 12; Length 312;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 IAKETNN 119  
 |||||  
 Db 244 IAKETNN 250

RESULT 56

Q9LD14  
ID Q9LD14 PRELIMINARY; PRT; 313 AA.  
AC Q9LD14;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Pinorensinol-lariciresinol reductase.  
OS Thuja plicata (Giant arborvitae).  
OC Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Thuja.  
OX NCBI\_TaxID=3316;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Stem;  
RX MEDLINE=99091594; PubMed=9872995;  
RA Fujita M., Gang D.R., Davin L.B., Lewis N.G.;  
RT "Recombinant pinorensinol-lariciresinol reductases from western red  
cedar (Thuja plicata) catalyze opposite enantiospecific conversions.";  
RL J. Biol. Chem. 274:618-627(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Stem;  
RX MEDLINE=99091594; PubMed=9872995;  
RA Fujita M., Gang D.R., Davin L.B., Lewis N.G.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF242503; AAF63507.1; -  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR InterPro; IPR003866; Isoflav reduct.  
DR Pfam; PF02716; Isoflavone reduct; 1.  
SQ SEQUENCE 313 AA; 35580 MW; 3D3178ACE37E8B37 CRC64;  
Query Match 4.9%; Score 7; DB 10; Length 313;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 130 KVRRAIE 136  
Db 140 KVRRAIE 146  
RESULT 57  
Q9LD00  
ID Q9LD00 PRELIMINARY; PRT; 314 AA.  
AC Q9LD00;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Phenylcoumaran benzylic ether reductase homolog Tpl (Pinorensinol-  
lariciresinol reductase).  
OS Thuja plicata (Giant arborvitae).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Thuja.  
OX NCBI\_TaxID=3316;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=YOUNG STEM;  
RX MEDLINE=99167523; PubMed=10066819;  
RA Gang D.R., Kasahara H., Xia Z.Q., Vander Mijnsbrugge K., Bauw G.,  
Boerjan W., Van Montagu M., Davin L.B., Lewis N.G.;  
RT "Evolution of plant defense mechanisms. Relationships of  
phenylcoumaran benzylic ether reductases to pinorensinol-lariciresinol  
and isoflavone reductases.";  
RL J. Biol. Chem. 274:7516-7527(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=YOUNG STEM;  
RX MEDLINE=99091594; PubMed=9872995;  
RA Gang D.R., Kasahara H., Xia Z.Q., Davin L.B., Lewis N.G.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=YOUNG STEM;  
RX MEDLINE=99091594; PubMed=9872995;  
RA Fujita M., Gang D.R., Davin L.B., Lewis N.G.;  
RT "Recombinant pinorensinol-lariciresinol reductases from western red

cedar (Thuja plicata) catalyze opposite enantiospecific conversions.";  
RL J. Biol. Chem. 274:618-627(1999).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=YOUNG STEM;  
RX Fujita M., Gang D.R., Davin L.B., Lewis N.G.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF242500; AAF64183.1; -  
DR EMBL; AF242505; AAF63509.1; -  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR InterPro; IPR003866; Isoflav reduct.  
DR Pfam; PF02716; Isoflavone reduct; 1.  
SQ SEQUENCE 314 AA; 35408 MW; 64BA4ED27EF6E8CB CRC64;  
Query Match 4.9%; Score 7; DB 10; Length 314;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 130 KVRRAIE 136  
Db 140 KVRRAIE 146  
RESULT 58  
OS9708  
ID OS9708 PRELIMINARY; PRT; 343 AA.  
AC OS9708;  
DT 01-AUG-1998 (TReMBLrel. 07, Created)  
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN SPEC36.11  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972h-;  
RA Lyne M., Wood V., Rajandream M.A., Barrell B.G., Hilbert H.,  
Moestl D., Duesterhoeft A;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL023589; CAA19059.1; -  
DR PIR; T40306; T40306  
DR GeneDB; Spombe; SPBC36.11; -  
KW Hypothetical protein.  
SQ SEQUENCE 343 AA; 37741 MW; A04C84F8C55DD5FF CRC64;  
Query Match 4.9%; Score 7; DB 3; Length 343;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 75 BEHKHS 81  
Db 330 BEHKHS 336  
RESULT 59  
Q7USB6  
ID Q7USB6 PRELIMINARY; PRT; 353 AA.  
AC Q7USB6;  
DT 01-OCT-2003 (TReMBLrel. 25, Created)  
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN R54603  
OS Rhodospirillum rubrum.  
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
OC Planctomycetaceae; Pirellula.  
OX NCBI\_TaxID=117;  
RN [1]  
RP SEQUENCE FROM N.A.

```

RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot I.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schleiner H., Anann R., Reinhardt R.,
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RL strain 1.",
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294140; CAD73881.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 353 AA; 39090 MW; 48EDB240A95B44C CRC64;

Query Match 4.9%; Score 7; DB 16; Length 353;
Best Local Similarity 100.0%; Pred. No. 95; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 136 EQLAAMD 142
Db 241 EQLAAMD 247

RESULT 50
Q9TXL0 PRELIMINARY; PRT; 374 AA.
AC Q9TXL0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN K04F1.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RC SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RA "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.",
RL Science 282:2012-2018(1998).
[2]
RC SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RA "The sequence of C. elegans cosmid K04F1.",
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
[3]
RC SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RX Waterston R.;
RA "Direct Submission.",
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF106575; AAC78165.1; -.
DR PIR; G88955; G88955.
DR WormPep; K04F1.6; CE19957.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR InterPro; IPR000494; EGFR_L domain.
DR Pfam; PF01030; Recep_L domain; I.
KW Hypothetical protein.
SQ SEQUENCE 374 AA; 43106 MW; 42BBE75C5ADBED10 CRC64;

Query Match 4.9%; Score 7; DB 5; Length 374;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 QFELTL 98
Db 236 QFELTL 242

```

```

RESULT 61
Q9RTR2 PRELIMINARY; PRT; 379 AA.
AC Q9RTR2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein DR1694.
GN DR1694.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
[1]
RC SEQUENCE FROM N.A.
RP STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.",
RL Science 286:1571-1577(1999).
DR EMBL; AE002011; ABF11251.1; -.
DR PIR; E75364; E75364.
DR TIGR; DR1694; -.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
DR InterPro; IPR000051; SAM bind.
KW Hypothetical protein; Complete proteome.
OX NCBI_TaxID=6239;
SQ SEQUENCE 379 AA; 41459 MW; 83459F87DEABDC4C CRC64;

Query Match 4.9%; Score 7; DB 16; Length 379;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 134 AIEQLAA 140
Db 262 AIEQLAA 268

RESULT 62
Q9I253 PRELIMINARY; PRT; 393 AA.
AC Q9I253;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Probable pyridoxal-phosphate dependent enzyme.
GN PA2062.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
[1]
RC SEQUENCE FROM N.A.
RP STRAIN=ATCC 15692 / PAOI;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Raizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen.",
RL Nature 406:959-964(2000).
DR EMBL; AE004633; AAG05450.1; -.
DR PIR; C83387; C83387.
DR GO; GO:0008483; F:transaminase activity; IEA.

```

DR GO: GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR000192; AminoTrans\_V.  
 DR Pfam; PF00266; aminotran\_5; 1.  
 DR PROSITE; PS00595; AA\_TRANSER\_CLASS\_5; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 393 AA; 41520 MW; 38099C516ECSEB5B CRC64;

Query Match 4.9%; Score 7; DB 16; Length 393;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PAWQPFLL 14  
 |||||  
 DB 381 PAWQPFLL 387

RESULT 63  
 Q9AVV0 PRELIMINARY; PRT; 405 AA.  
 AC Q8AVV0;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to splicing factor 3a, subunit 2, 66kDa.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.; to the EMBL/GenBank/DBJ databases.  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC041254; AA41254.1; .  
 DR GO: GO:0005634; C:nucleus; IEA.  
 DR GO: GO:0003723; F:RNA binding; IEA.  
 DR InterPro; IPR000690; Znf\_matriin.  
 DR InterPro; IPR003604; Znf\_U1.  
 DR SMART; SMO0451; ZNF\_U1.  
 DR PROSITE; PSS0171; ZF\_MATRIIN; 1.  
 SQ SEQUENCE 405 AA; 43673 MW; 2D91D3FCF2BFD6A0 CRC64;

Query Match 4.9%; Score 7; DB 13; Length 405;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPPA 9  
 |||||  
 DB 220 APTLPPA 226

RESULT 64  
 Q9SCQ5 PRELIMINARY; PRT; 427 AA.  
 AC Q9SCQ5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN T3A5.90.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bloembergen H., Mewes H.W., Lemcke K., Mayer K.F.X., Quetier F.,  
 RA Salanoubat M.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL132979; CAB62440.1; .  
 DR PIR; T46148; T46148.  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR006566; FBD.  
 DR InterPro; IPR001092; HLH\_basic.  
 DR InterPro; IPR008945; SKP1\_skp2.  
 DR Pfam; PF00646; F-box; 1.  
 DR SMART; SMO0579; FBD; 1.  
 DR SMART; SMO0256; FBOX; 1.  
 DR PROSITE; PS00038; HLH\_1; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 427 AA; 49869 MW; DD61F344D0D2C003 CRC64;

Query Match 4.9%; Score 7; DB 10; Length 427;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 DDPIEEH 77  
 |||||  
 DB 339 DDPIEEH 345

RESULT 65  
 Q26582 PRELIMINARY; PRT; 442 AA.  
 AC Q26582;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Heat shock protein 86 (fragment).  
 OS Schistosoma mansoni (Blood fluke).  
 OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;  
 OC Schistosomatidae; Schistosomatidae; Schistosoma.  
 OX NCBI\_TaxID=6183;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90043865; PubMed=2509907;  
 RA Johnson K.S., Wells K., Bock J.V., Nene V., Taylor D.W.,  
 RA Cordingley J.S.;  
 RT "the 86 kDa antigen from Schistosoma mansoni is a heat shock protein  
 homologous to yeast HSP-90.";  
 RL Mol. Biochem. Parasitol. 36:19-28(1989).  
 DR EMBL; J04017; AAA29899.1; .  
 DR PIR; A45529; A45529.  
 DR GO: GO:0003754; F:chaperone activity; IEA.  
 DR GO: GO:0004289; F:subtilase activity; IEA.  
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001404; Hsp90.  
 DR InterPro; IPR000209; Peptidase\_S8.  
 DR Pfam; PF00183; HSP90; 1.  
 DR PROSITE; PS00136; SUBTILASE\_ASP; 1.  
 FT NON\_TER  
 SQ SEQUENCE 442 AA; 50996 MW; 48B83001936D6DBB CRC64;

Query Match 4.9%; Score 7; DB 5; Length 442;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 KQOFEEL 96  
 |||||  
 DB 269 KQOFEEL 275

RESULT 66  
 Q89ZX8 PRELIMINARY; PRT; 467 AA.  
 AC Q89ZX8;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative oxidoreductase (Putative secreted protein).

```
GN BT4243.
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550859; PubMed=12663928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RA "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AE016944; AA079348.1; -.
DR GO; GO:0016431; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000683; GFO_IDH_MoCA.
DR Pfam; PF01408; GFO_IDH_MoCA; 1.
KW Complete proteome.
SQ SEQUENCE 467 AA; 52831 MW; E0CE1286819FC2FB CRC64;

Query Match 4.9%; Score 7; DB 16; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 125 EETAKV 131
Db 387 EETAKV 393

RESULT 67
Q8F9W0 PRELIMINARY; PRT; 501 AA.
ID Q8F9W0;
AC Q8F9W0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN LA0078.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011199; AA047277.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 501 AA; 57403 MW; 46A887E1E8AA34DD CRC64;

Query Match 4.9%; Score 7; DB 16; Length 501;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 EFLKLD 106
Db 185 EFLKLD 191

RESULT 68
Q97E19 PRELIMINARY; PRT; 525 AA.
ID Q97E19;
AC Q97E19;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Spore germination protein GerKA, membrane protein.
GN CAC3300.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;

[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007827; AA081233.1; -.
DR PIR; F97305; F97305.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0009847; P:spore germination; IEA.
DR InterPro; IPR004995; GerA.
DR Pfam; PF03323; GerA; 1.
KW Complete proteome.
SQ SEQUENCE 525 AA; 58570 MW; BA2E14BF1818C4EA CRC64;

Query Match 4.9%; Score 7; DB 16; Length 525;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 BELTGE 100
Db 183 BELTGE 189

RESULT 69
Q87YV7 PRELIMINARY; PRT; 546 AA.
ID Q87YV7;
AC Q87YV7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Methyl-accepting chemotaxis protein.
GN PSPT03685.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
RA White O., Fraser C., Collier A.;
RA "Complete sequence of Pseudomonas syringae.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016869; AA057154.1; -.
DR TIGR; PSPT03685; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004871; P:signal transducer activity; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0007165; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMF.
DR InterPro; IPR000005; HTHarac.
DR InterPro; IPR004090; Me_chemotaxis.
DR Pfam; PF00672; HAMF; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDUCR.
DR PROSITE; PS01111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
KW Complete proteome.
SQ SEQUENCE 546 AA; 58319 MW; CD58FF6BE35AC1A0 CRC64;
```

Query Match 4.9%; Score 7; DB 16; Length 546;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 AIEQLAA 140  
Db 350 AIEQLAA 356

RESULT 70  
Q9LD30 PRELIMINARY; PRT; 579 AA.  
AC Q9LD30;  
DT 01-OCT-2000 (TREMELrel. 15, Created)  
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)  
DE 01-OCT-2000 (TREMELrel. 15, Last annotation update)  
DE Dipl-associated protein G (fragment).  
GN DAPG.  
OS Cryptocodinium cohnii (Dinoflagellate).  
OC Eukaryota; Alveolata; Dinophyceae; Gonyaulacales; Cryptocodiniaceae;  
OC Cryptocodinium.  
OX NCBI\_TaxID=2866;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9305750; PubMed=10377987;  
RA Bhaid Y., Gerard M.L., Ausell J., Soyer-Gobillard M.O., Moreau H.;  
RT "Cyclic expression of a nuclear protein in a dinoflagellate.";  
RN J. Eukaryot. Microbiol. 46:259-267(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Gullebaud D., Moreau H.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF255447; AAF78929.1; -;  
FT NON TER 1  
FT NON TER 579  
SQ SEQUENCE 579 AA; 65085 MW; FOC05B3CE5AE58AE CRC64;

Query Match 4.9%; Score 7; DB 10; Length 579;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PTLPPAW 10  
Db 160 PTLPPAW 166

RESULT 71  
Q87FL8 PRELIMINARY; PRT; 609 AA.  
AC Q87FL8;  
DT 01-JUN-2003 (TREMELrel. 24, Created)  
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)  
DE 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Putative AcetD.  
GN VPA1661.  
OS Vibrio parahaemolyticus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=670;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RIMD 2210633 / Serotype O3:k6;  
RX MEDLINE=22508454; PubMed=12620739;  
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,  
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
RA Yasunaga T., Honda T., Shingawa H., Hattori M., Iida T.;  
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism distinct from that of V. cholerae.";  
RL Lancet 361:743-749(2003).  
DR EMBL; AF005089; BAC63004.1; -;  
DR GO; GO:0015343; F:siderochrome-iron transporter activity; IEA.  
DR GO; GO:0019290; P:siderochrome biosynthesis; IEA.  
DR InterPro; IPR007310; IuckA\_lucc.

DR Pfam; PF04183; IuckA\_lucC; 1.  
KW Complete proteome.  
SQ SEQUENCE 609 AA; 69486 MW; 148CA003AAF2099E CRC64;

Query Match 4.9%; Score 7; DB 16; Length 609;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 VRAAIEQ 137  
Db 262 VRAAIEQ 268

RESULT 72  
Q8PZW5 PRELIMINARY; PRT; 625 AA.  
AC Q8PZW5;  
DT 01-OCT-2002 (TREMELrel. 22, Created)  
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)  
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)  
DE Conserved protein.  
GN MW03774.  
OS Methanosarcina mazei (Methanosarcina frisia).  
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;  
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.  
OX NCBI\_TaxID=2209;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Goel / GO1 / ATCC BAA-199 / DSM 3647 / OCM 88;  
RX MEDLINE=22120827; PubMed=12125824;  
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,  
RA Martinez-Arias R., Henne A., Wierzer A., Baumer S., Jacobi C.,  
RA Bruggemann H., Lienard T., Christmann A., Soenneke M., Steckel S.,  
RA Bhattacharya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,  
RA Fritz H.-J., Gottschalk G.;  
RT "The genome of Methanosarcina mazei: evidence for lateral gene transfer between Bacteria and Archaea.";  
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).  
DR EMBL; AE013262; AAM30070.1; -;  
KW Complete proteome.  
SQ SEQUENCE 625 AA; 68028 MW; 811CFB57502DA3A1 CRC64;

Query Match 4.9%; Score 7; DB 17; Length 625;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 LKLDRE 108  
Db 40 LKLDRE 46

RESULT 73  
Q8EUJ0 PRELIMINARY; PRT; 671 AA.  
AC Q8EUJ0;  
DT 01-MAR-2003 (TREMELrel. 23, Created)  
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)  
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DE Phosphate ABC transporter permease PstA.  
GN MYPE9360.  
OS Mycoplasma penetrans.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=28227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HF-2;  
RX MEDLINE=22354719; PubMed=12466555;  
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,  
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;  
RT "The complete genomic sequence of Mycoplasma penetrans, an intracellular bacterial pathogen in humans.";  
RL Nucleic Acids Res. 30:5293-5300(2002).  
DR EMBL; AP004174; BAC44723.1; -;



DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000315; BPD transp.  
 DR Pfam; PF00528; BPD transp; 2.  
 KW Complete proteome.  
 SQ SEQUENCE 671 AA; 74044 MW; BBFC521F4CC311A3 CRC64;

Query Match 4.9%; Score 7; DB 16; Length 671;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 AKNKIAK 115  
 Db 452 AKNKIAK 458  
 |||||

RESULT 74  
 Q91VK6 PRELIMINARY; PRT; 673 AA.

AC Q91VK6  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast tumor;  
 RA Strausberg R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC012871; AAH12871.1;  
 DR InterPro; IPR004018; RPEL\_repeat.  
 DR Pfam; PF02755; RPEL; 4.  
 DR SMART; SM00707; RPEL; 4.  
 KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 673 AA; 73270 MW; 3D938DC6A2505820 CRC64;

Query Match 4.9%; Score 7; DB 11; Length 673;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 ERMAEAG 42  
 Db 46 ERMAEAG 52  
 |||||

RESULT 75  
 Q8ZJ11 PRELIMINARY; PRT; 698 AA.

AC Q8ZJ11  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE 4-alpha-glucanotransferase (EC 2.4.1.25) (Amylomaltase).  
 GN MALQ OR YPO0126 OR Y3902.  
 OS Versinia pestis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Yersinia.  
 OX NCBI\_TaxID=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CO-92 / Biovar Orientalis;  
 RX MEDLINE=21470413; PubMed=11596350;  
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,  
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Parraga A.M.,  
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,

RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;  
 RT "Genome sequence of Yersinia pestis, the causative agent of plague."  
 RL Nature 413:523-527(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KIM5 / Biovar Mediaevalis;  
 RX MEDLINE=22137863; PubMed=12142430;  
 RA Dang W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
 RA Straley S.C., McDonough K.A., Milles M.L., Matson J.S., Blattner F.R.,  
 RA Perry R.D.;  
 RT "Genome sequence of Yersinia pestis KIM."  
 RL J. Bacteriol. 184:4601-4611(2002).  
 DR EMBL; AJ414141; CAC89389.1; -;  
 DR EMBL; AE013995; AM87446.1; -;  
 DR PIR; AC0016; AC0016.  
 DR GO; GO:0004134; F:4-alpha-glucanotransferase activity; IEA.  
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR003385; Glyco\_hydro\_77.  
 DR Pfam; PF02446; 4A\_glucoamylans; 1.  
 DR TIGRfams; TIGR00217; malQ.1.  
 KW Glycosyltransferase; Transferase; Complete proteome.  
 SQ SEQUENCE 698 AA; 78553 MW; 685512C27B2A682 CRC64;

Query Match 4.9%; Score 7; DB 16; Length 698;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 LDRERAK 110  
 Db 586 LDRERAK 592  
 |||||

Search completed: August 11, 2004, 14:25:06  
 Job time : 42 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 14:17:41 ; Search time 54 Seconds  
(without alignments)  
742.995 Million cell updates/sec

Title: US-09-690-825-34  
Perfect score: 142  
Sequence: 1 MGATLPPAWOPFLKDHRS.....EFETAKVRRRAIFQLAAMD 142

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : A\_Geneseq\_29Jan04: \*  
1: Geneseq1980s: \*  
2: Geneseq1990s: \*  
3: Geneseq2000s: \*  
4: Geneseq2001s: \*  
5: Geneseq2002s: \*  
6: Geneseq2003as: \*  
7: Geneseq2003bs: \*  
8: Geneseq2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	142	100.0	142	2	AAM61359 Survivin
2	142	100.0	142	2	AAY49080 Human Sur
3	142	100.0	142	4	AAG65360 Human Her
4	142	100.0	142	5	Aau74571 Human sur
5	142	100.0	142	5	Aao14948 Survivin-
6	142	100.0	142	6	Abp72162 Human Sur
7	142	100.0	142	7	Abp72162 Human Sur
8	142	100.0	143	4	Abp72162 Human Sur
9	132	93.0	142	5	Aau74573 Human mut
10	129	90.8	141	6	Abu56477 Lung can
11	125	88.0	125	5	Aau74582 Human del
12	108	76.1	142	5	Aau74574 Human mut
13	96	67.6	142	5	Aau74577 Human mut
14	88	62.0	142	5	Aau74572 Human mut
15	87	61.3	142	5	Aau74578 Human mut
16	87	61.3	142	5	Aau74579 Human mut
17	82	57.7	142	5	Aau74580 Human mut
18	82	57.7	142	5	Aau74581 Human mut
19	79	55.6	142	5	Aau74575 Human mut
20	76	53.5	116	3	AGG02311 Human sec
21	75	52.8	75	7	ADB1824 Human inh
22	75	52.8	142	5	Aau74576 Human sur
23	74	52.1	165	6	Ada09936 Human sur
24	59	41.5	83	7	ADB1823 Human inh
25	41	28.9	140	2	Aaw19749 Mouse inh

26	41	28.9	140	5	Aau74583	Aau74583 Mouse sur
27	36	25.4	155	4	Aau30231	Aau30231 Novel hum
28	24	16.9	24	6	ABP70506	ABP70506 Peptide c
29	21	14.8	21	6	ABP70505	ABP70505 Peptide c
30	20	14.1	20	2	AAM61358	AAM61358 Survivin
31	17	12.0	17	2	AAM61360	AAM61360 Survivin
32	17	12.0	17	2	AAE11066	Aae11066 Human sur
33	10	7.0	10	6	ADA09938	Ada09938 Human HLA
34	10	7.0	17	4	AAE11070	AAe11070 Human sur
35	9	6.3	9	4	AAE11069	AAe11069 Human sur
36	9	6.3	1822	6	ADA09937	Ada09937 Human HLA
37	8	5.6	9	5	ABU57366	Abu57366 Surl/M2 p
38	7	4.9	9	5	ABU57366	Abu57366 Peptide #
39	7	4.9	73	4	ABM19295	ABm19295 Peptide #
40	7	4.9	73	4	ABM38577	ABb38577 Peptide #
41	7	4.9	73	4	ABM32028	ABm32028 Peptide #
42	7	4.9	73	4	ABM32028	ABb32028 Peptide #
43	7	4.9	73	4	ABM32028	ABb32028 Peptide #
44	7	4.9	73	4	ABM32028	ABb32028 Peptide #
45	7	4.9	73	4	ABM32028	ABb32028 Peptide #
46	7	4.9	73	4	ABM32028	ABb32028 Peptide #
47	7	4.9	73	4	ABM32028	ABb32028 Peptide #
48	7	4.9	73	4	ABM32028	ABb32028 Peptide #
49	7	4.9	73	4	ABM32028	ABb32028 Peptide #
50	7	4.9	73	4	ABM32028	ABb32028 Peptide #
51	7	4.9	73	4	ABM32028	ABb32028 Peptide #
52	7	4.9	73	4	ABM32028	ABb32028 Peptide #
53	7	4.9	73	4	ABM32028	ABb32028 Peptide #
54	7	4.9	73	4	ABM32028	ABb32028 Peptide #
55	7	4.9	73	4	ABM32028	ABb32028 Peptide #
56	7	4.9	73	4	ABM32028	ABb32028 Peptide #
57	7	4.9	73	4	ABM32028	ABb32028 Peptide #
58	7	4.9	73	4	ABM32028	ABb32028 Peptide #
59	7	4.9	73	4	ABM32028	ABb32028 Peptide #
60	7	4.9	73	4	ABM32028	ABb32028 Peptide #
61	7	4.9	73	4	ABM32028	ABb32028 Peptide #
62	7	4.9	73	4	ABM32028	ABb32028 Peptide #
63	7	4.9	73	4	ABM32028	ABb32028 Peptide #
64	7	4.9	73	4	ABM32028	ABb32028 Peptide #
65	7	4.9	73	4	ABM32028	ABb32028 Peptide #
66	7	4.9	73	4	ABM32028	ABb32028 Peptide #
67	7	4.9	73	4	ABM32028	ABb32028 Peptide #
68	7	4.9	73	4	ABM32028	ABb32028 Peptide #
69	7	4.9	73	4	ABM32028	ABb32028 Peptide #
70	7	4.9	73	4	ABM32028	ABb32028 Peptide #
71	7	4.9	73	4	ABM32028	ABb32028 Peptide #
72	7	4.9	73	4	ABM32028	ABb32028 Peptide #
73	7	4.9	73	4	ABM32028	ABb32028 Peptide #
74	7	4.9	73	4	ABM32028	ABb32028 Peptide #
75	7	4.9	73	4	ABM32028	ABb32028 Peptide #
76	7	4.9	73	4	ABM32028	ABb32028 Peptide #
77	7	4.9	73	4	ABM32028	ABb32028 Peptide #
78	6	4.2	11	2	AAM95149	Aaw95149 Bacillus
79	6	4.2	11	2	AAM95149	Aaw95149 Bacillus
80	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
81	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
82	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
83	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
84	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
85	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
86	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
87	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
88	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
89	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
90	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
91	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
92	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
93	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
94	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
95	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
96	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
97	6	4.2	12	2	AAM95149	Aaw95149 Bacillus
98	6	4.2	12	2	AAM95149	Aaw95149 Bacillus

99 6 4.2 65 4 AAM43557 Human pol  
100 6 4.2 65 4 AAU19925 Novel hum

## ALIGNMENTS

## RESULT 1

AAW61359  
ID AAW61359 standard; protein; 142 AA.

XX  
AC AAW61359;

XX 25-SEP-1998 (first entry)

XX Survivin protein.

XX survivin; apoptosis; cellular apoptosis; transplantation;

KW motor neuron degenerative disease; HIV infection; immunosuppression;

KW gastrointestinal perturbations; cardiovascular disorder.

XX Homo sapiens.

XX WO9822589-A2.

XX 28-MAY-1998.

XX 20-NOV-1997; 97WO-US021880.

XX 20-NOV-1996; 96US-0031435P.

XX 20-NOV-1997; 97US-00975080.

XX (UYVA ) UNIV YALE.

XX Altieri DC;

XX WPI; 1998-312475/27.

XX N-PSDB; AAV27941.

XX Modulating apoptosis by controlling the Survivin gene - useful for

XX treating transplant rejection, degenerative disorders and tumours.

XX Disclosure; Fig 10; 108pp; English.

XX The survivin gene can be used to control apoptosis through modification

XX of the gene. Survivin peptides can be used to inhibit cellular apoptosis,

XX e.g. for enhancing the viability of organs and tissues prior to their

XX transplantation, for preserving the growth of cells in culture or for

XX treating conditions involving abnormal apoptosis, e.g. degenerative

XX diseases such as motor neuron degenerative diseases, HIV infection,

XX dermatological effects of ageing, disorders and diseases such as

XX immunosuppression, gastrointestinal perturbations, cardiovascular

XX disorders, apoptosis related to reperfusion damage, rejection of tissue

XX transplantation and Alzheimer's disease. Agents which block Survivin

XX activity can be used to treat e.g. tumours

XX Sequence 142 AA;

XX Query Match 100.0%; Score 142; DB 2; Length 142;

XX Best Local Similarity 100.0%; Pred. No. 1e-142;

XX Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 2

AA49080  
ID AAY49080 standard; protein; 142 AA.

XX  
AC AAY49080;

XX 17-JAN-2000 (first entry)

XX Human Survivin amino acid sequence.

XX Survivin; IAP; apoptosis inhibitor; tumour cell; mitotic spindle;

KW effector cell protease receptor-1; EPR-1; embryonic tissue; tubulin;

KW abnormal cell growth; tumorigenesis.

XX Homo sapiens.

XX WO9950440-A2.

XX 07-OCT-1999.

XX 01-APR-1999; 99WO-US007205.

XX 01-APR-1998; 98US-0080288P.

XX (UYVA ) UNIV YALE.

XX Altieri DC;

XX WPI; 1999-591329/50.

XX N-PSDB; AAZ31037.

XX Identifying agents that modulate Survivin interactions.

XX Example; Fig 8A; 56pp; English.

XX This is the amino acid sequence of the human survivin protein. Survivin

XX is a 142 amino acid protein of approximately 16.5kD. Survivin is a member

XX of the IAP family of apoptosis inhibitors, and the gene is located on

XX chromosome 17q25. The nucleic acid sequence of Survivin is related to the

XX Effector Cell Protease Receptor-1 (EPR-1). Survivin is expressed in

XX tumour cells and embryonic tissue. The sequence is used in the invention

XX which involves identifying an agent that modulates interactions between

XX Survivin and tubulin. Agents that modulate interaction between Survivin

XX and tubulin or mitotic spindles may be administered to a mammal to

XX modulate biological or pathological processes mediated by Survivin, for

XX example Survivin-mediated inhibition of cellular apoptosis. Thus an

XX inhibiting agent may be used to block abnormal cell growth, for example

XX during tumorigenesis. An agent that increases interaction may be used to

XX extend cell growth in culture

XX Sequence 142 AA;

XX Query Match 100.0%; Score 142; DB 2; Length 142;

XX Best Local Similarity 100.0%; Pred. No. 1e-142;

XX Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGATLPPAWQFLKDHRISTFKNWPFLGCACTPERMAEAGFIHCPTNEPDLAQCF 60

Db 1 MGATLPPAWQFLKDHRISTFKNWPFLGCACTPERMAEAGFIHCPTNEPDLAQCF 60

QY 61 FKELEGNEPDDPTEEHKHSAGCAFLSVKKQFPELTGLGFLKDLDRANKKIATETNNK 120

Db 61 FKELEGNEPDDPTEEHKHSAGCAFLSVKKQFPELTGLGFLKDLDRANKKIATETNNK 120

QY 121 KKEFEETAKKVRRAIEQLAAMD 142

Db 121 KKEFEETAKKVRRAIEQLAAMD 142

## RESULT 3

AAG65360

ID AC AAG65360 standard; protein; 142 AA.  
 XX AC AAG65360;  
 XX DT 30-NOV-2001 (first entry)  
 XX DE Human Her-3 polypeptide sequence.  
 XX KW Her-3; epidermal growth factor; EGF; receptor/tyrosine kinase; human;  
 XX KW antinflammatory; cytostatic; antibacterial; antisense.  
 XX OS Homo sapiens.  
 XX PN US6277640-B1.  
 XX PD 21-AUG-2001.  
 XX PF 31-JUL-2000; 2000US-00630706.  
 XX PR 31-JUL-2000; 2000US-00630706.  
 XX PA (ISIS-) ISIS PHARM INC.  
 XX PI Bennett CF, Cowser LM;  
 XX DR WPI; 2001-535134/S9.  
 XX DR N-PSDB; AA47531.  
 XX PT Antisense compounds capable of modulating expression of human Her-3,  
 PT member of epidermal growth factor family of receptor/tyrosine kinases,  
 PT useful for preventing or delaying infection, inflammation or tumor  
 PT formation.  
 XX PS Example 15; Col 57-70; 49pp; English.  
 XX CC The invention provides antisense compounds capable of inhibiting the  
 CC expression of human Her-3, a member of epidermal growth factor (EGF)  
 CC family of receptor/tyrosine kinases. The antisense oligonucleotides are  
 CC useful for inhibiting the expression of Her-3 in cells or tissues. They  
 CC are commonly used as research reagents and in diagnostics for example, to  
 CC elucidate the function of particular genes. The antisense compounds are  
 CC also useful for distinguishing between functions of various members of a  
 CC biological pathway and for research use. They are also utilized for  
 CC diagnostics, therapeutics, prophylaxis and in kits. They are useful  
 CC prophylactically, e.g. to prevent or delay infection, inflammation or  
 CC tumor formation. The present sequence represents the human Her-3 sequence  
 CC (GenBank Accession No. U75285)  
 XX SQ Sequence 142 AA;  
 Query Match 100.0%; Score 142; DB 4; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 1e-142;  
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGAPTLPPAQPFLKDHRISTFKWPFLEGGCACTPERMAEAGFIHCPTEHPDLAQCFPC 60  
 Db 1 MGAPTLPPAQPFLKDHRISTFKWPFLEGGCACTPERMAEAGFIHCPTEHPDLAQCFPC 60  
 QY 61 FKELEGWEPDDPIEBHKHSSGCFLSVKQKQFEELTLGEFLKDRERAKNKIAKETNNK 120  
 Db 61 FKELEGWEPDDPIEBHKHSSGCFLSVKQKQFEELTLGEFLKDRERAKNKIAKETNNK 120  
 QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 Db 121 KKEFEETAKKVRRAIEQLAAMD 142  
 RESULT 4  
 ID AAU74571  
 XX AAU74571 standard; protein; 142 AA.  
 XX AC AAU74571;  
 XX

DT 08-MAY-2002 (first entry)  
 XX DE Human survivin polypeptide.  
 XX KW Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
 KW tubulin interaction; caspase-3; CBK4; zinc chelation activity; apoptosis;  
 KW dimerisation; ovary; breast; pancreas; central nervous system; blood;  
 KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
 KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
 KW cell proliferative disorder; acquired immunodeficiency syndrome;  
 KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
 KW myelodysplastic syndrome; protein coordinate data.  
 XX OS Homo sapiens.  
 XX PN WC200202622-A2.  
 XX PD 10-JAN-2002.  
 XX PF 29-JUN-2001; 2001WO-US020872.  
 XX PR 29-JUN-2000; 2000US-00608352.  
 XX PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX PI Noel JP, Verdacia M, Hunter T, Huang H;  
 XX DR WPI; 2002-171638/22.  
 XX PT Novel isolated crystalline survivin polypeptide, useful to screen and  
 PT design compounds that bind to or interact with inhibitor of apoptosis  
 PT protein and protein family members, and for design of novel therapeutics.  
 XX PS Claim 42; Fig 2; 102pp; English.  
 XX CC The invention relates to an isolated crystalline human survivin  
 CC polypeptide, an inhibitor of apoptosis protein (IAP). The polypeptide is  
 CC useful for determining the ability of a survivin binding agent to  
 CC modulate tubulin interaction, p21 caspase-3 and CDK4 requirement or zinc  
 CC chelation activity and for identifying an agent such as a peptide or zinc  
 CC peptidomimetic which inhibits dimerisation of survivin. The ability of an  
 CC agent to modulate dimerisation can be determined through detection of a  
 CC change in apoptosis in a target cell expressing survivin. Binding agents  
 CC are useful for increasing apoptosis in a cell derived from a tissue  
 CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,  
 CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous  
 CC system, prostate, colon, rectum, cervix or endometrium, with a cell  
 CC proliferative disorder such as cancer. The molecules can also be used in  
 CC treatment or prevention of apoptosis which occurs as a part of AIDS,  
 CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease  
 CC and myelodysplastic syndromes. This sequence represents the human  
 CC survivin polypeptide  
 XX SQ Sequence 142 AA;  
 Query Match 100.0%; Score 142; DB 5; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 1e-142;  
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGAPTLPPAQPFLKDHRISTFKWPFLEGGCACTPERMAEAGFIHCPTEHPDLAQCFPC 60  
 Db 1 MGAPTLPPAQPFLKDHRISTFKWPFLEGGCACTPERMAEAGFIHCPTEHPDLAQCFPC 60  
 QY 61 FKELEGWEPDDPIEBHKHSSGCFLSVKQKQFEELTLGEFLKDRERAKNKIAKETNNK 120  
 Db 61 FKELEGWEPDDPIEBHKHSSGCFLSVKQKQFEELTLGEFLKDRERAKNKIAKETNNK 120  
 QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 Db 121 KKEFEETAKKVRRAIEQLAAMD 142  
 RESULT 5

AAO14948  
ID AAO14948 standard; protein; 142 AA.  
XX  
AC AAO14948;  
XX  
DT 06-AUG-2002 (first entry)  
XX  
DE Survivin-like protein 6.  
XX  
KW Survivin-like protein; diagnosis; screening; cancer;  
KW apoptosis abnormality; gene therapy.  
XX  
OS Unidentified.  
XX  
PN WO200233071-A1.  
XX  
PD 25-APR-2002.  
XX  
PF 16-OCT-2001; 2001WO-JP009071.  
XX  
PR 17-OCT-2000; 2000JP-00316721.  
PR 20-DEC-2000; 2000JP-00386809.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Tanaka H, Kaieda I;  
XX  
DR WPI; 2002-435536/46.  
DR N-PSDB; AAL42859.  
XX  
PT Baculovirus IAP repeat domain or RING-finger domain-containing survivin-like polypeptides and encoded DNAs, applicable in diagnosis and screening compounds for treating various cancers and apoptosis abnormality.  
XX  
PS Disclosure; Page 122-123; 136pp; Japanese.  
XX  
SQ The invention comprises the amino acid and coding sequences of survivin-like proteins. The survivin-like DNA and protein sequences are useful in diagnostics and screening compounds for treating various cancers and apoptosis abnormality, including gene therapy. The present amino acid sequence represents a survivin-like protein of the invention  
XX  
SQ Sequence 142 AA;  
Query Match 100.0%; Score 142; DB 5; Length 142;  
Best Local Similarity 100.0%; Pred. No. 1e-142;  
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGAPTLPPAQPFKDHRISTFKNWPFLGCACTPERMAEAGFIHCPTNEPDLAQCFPC 60  
Db 1 MGAPTLPPAQPFKDHRISTFKNWPFLGCACTPERMAEAGFIHCPTNEPDLAQCFPC 60  
QY 61 FKELEGWEPDDPIEBHKHSSGCAFLSVKKQFEELTLGFEFLKDRERAKNKIAKETNNK 120  
Db 61 FKELEGWEPDDPIEBHKHSSGCAFLSVKKQFEELTLGFEFLKDRERAKNKIAKETNNK 120  
QY 121 KKEFEETAKVRRRAIEQLAAMD 142  
Db 121 KKEFEETAKVRRRAIEQLAAMD 142  
RESULT 6  
ABP72162  
ID ABP72162 standard; protein; 142 AA.  
XX  
AC ABP72162;  
XX  
DT 22-APR-2003 (first entry)  
XX  
DE Human Survivin.  
XX  
KW Human; Survivin; inhibitor of apoptosis; cell death; apoptosis; cancer; cytostatic; cardiant; neuroprotective; gene therapy.

XX Homo sapiens.  
XX WO2003004606-A2.  
XX 16-JAN-2003.  
XX 03-JUL-2002; 2002WO-US021002.  
XX 03-JUL-2001; 2001US-00898158.  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX Troy CM, Shelanski ML;  
XX WPI; 2003-210351/20.  
XX N-PSDB; ABZ58106.  
XX New nucleic acid encoding an inhibitor-of-apoptosis protein, useful for treating cancer, neurodegenerative disorder or cardiomyopathy.  
XX Disclosure; Fig 20A; 124pp; English.  
XX The present sequence is the protein sequence of human Survivin protein. The invention provides a nucleic acid, such as an antisense oligonucleotide, which specifically hybridizes to a nucleic acid encoding an inhibitor of apoptosis protein. Survivin is an example of an inhibitor of apoptosis or inducer of cell death protein of the invention. A claimed method for inducing a cell's death comprises contacting the cell with the nucleic acid under conditions permitting the nucleic acid to enter the cell, especially the use of a vector, liposome, or a mechanical or electrical means. The method is used to treat acute lymphocytic leukaemia, acute myelogenous leukaemia, lung cancer, breast cancer, ovarian cancer, prostate cancer, lymphoma, Hodgkin's disease, malignant melanoma, neuroblastoma, renal cell carcinoma and squamous cell carcinoma (all claimed). The invention also provides a second nucleic acid, which specifically hybridizes to a nucleic acid encoding a protein, other than caspase-2, that induces cell death. A claimed method for inhibiting a cell's death comprises contacting the cell with the nucleic acid under conditions permitting the nucleic acid to enter the cell. The method is used to treat a neurodegenerative disorder (especially a brain disorder or central nervous system disorder), or a heart disorder (especially cardiomyopathy) in a human (all claimed)  
XX  
SQ Sequence 142 AA;  
Query Match 100.0%; Score 142; DB 6; Length 142;  
Best Local Similarity 100.0%; Pred. No. 1e-142;  
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGAPTLPPAQPFKDHRISTFKNWPFLGCACTPERMAEAGFIHCPTNEPDLAQCFPC 60  
Db 1 MGAPTLPPAQPFKDHRISTFKNWPFLGCACTPERMAEAGFIHCPTNEPDLAQCFPC 60  
QY 61 FKELEGWEPDDPIEBHKHSSGCAFLSVKKQFEELTLGFEFLKDRERAKNKIAKETNNK 120  
Db 61 FKELEGWEPDDPIEBHKHSSGCAFLSVKKQFEELTLGFEFLKDRERAKNKIAKETNNK 120  
QY 121 KKEFEETAKVRRRAIEQLAAMD 142  
Db 121 KKEFEETAKVRRRAIEQLAAMD 142  
RESULT 7  
ABR61582  
ID ABR61582 standard; protein; 142 AA.  
XX  
AC ABR61582;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Human survivin protein.  
XX

KW RasGAP; Aurora kinase; cytostatic; anti-tumoural; human; survivin.  
 XX Homo sapiens.  
 OS W02003087395-A2.  
 PN 23-OCT-2003.  
 XX 15-APR-2003; 2003WO-15002972.  
 XX 15-APR-2002; 2002US-0372483P.  
 XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 XX Garbay C, Gigoux V, Camonis J, L'hoste S, Samson J;  
 XX WPI; 2003-845340/78.  
 XX N-PSDB; ACF58071.  
 XX Identifying anti-tumoral compounds, comprises determining the capacity of  
 PT a compound to inhibit interaction between RasGAP and Drosophila  
 PT melanogaster (Dm) aurora kinase, or between RasGAP, Dm aurora kinase and  
 PT aurora binding protein.  
 XX Disclosure; Page 67; Opp; English.  
 XX The invention relates to identifying a biologically active compound with  
 CC anti-tumoural properties, where the compound is studied for its capacity  
 CC to inhibit the interaction between: (a) RasGAP and the Drosophila  
 CC melanogaster Aurora kinase or an orthologue of the kinase, or its  
 CC fragment able to interact with RasGAP; or (b) RasGAP, an Aurora-binding  
 CC protein and the D. melanogaster Aurora kinase or an orthologue of the  
 CC kinase, or its fragment. The method is useful for identifying anti-  
 CC tumoural agents. The present sequence represents a human survivin  
 CC protein, an aurora-binding protein  
 XX  
 XX Sequence 142 AA;  
 SQ  
 Query Match 100.0%; Score 142; DB 7; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 1e-142;  
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGAPTLPPAWQPLKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60  
 DB 1 MGAPTLPPAWQPLKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60  
 QY 61 FKELEGWEPDDPIEBHKHSSGCAFLSVKKQFEBLTGFEFLKDRERAKNKIAKETNNK 120  
 DB 61 FKELEGWEPDDPIEBHKHSSGCAFLSVKKQFEBLTGFEFLKDRERAKNKIAKETNNK 120  
 QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 DB 121 KKEFEETAKKVRRAIEQLAAMD 142  
 RESULT 8  
 AAU29999  
 ID AAU29999 standard; protein; 143 AA.  
 XX AAU29999;  
 AC  
 XX 18-DEC-2001 (first entry)  
 DT  
 XX Novel human secreted protein #490.  
 DE  
 XX Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX Homo sapiens.  
 OS W0200179449-A2.  
 PN  
 XX

PD 25-OCT-2001.  
 XX 16-APR-2001; 2001WO-US008636.  
 XX 18-APR-2000; 2000US-00552929.  
 PR 26-JAN-2001; 2001US-00770160.  
 XX (HYSB-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 PI WPI; 2001-611725/70.  
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.  
 XX Claim 20; Page 219; 765pp; English.  
 XX The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX  
 XX Sequence 143 AA;  
 SQ  
 Query Match 100.0%; Score 142; DB 4; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 1e-142;  
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGAPTLPPAWQPLKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60  
 DB 1 MGAPTLPPAWQPLKDHRISTFKNWPFLGGCACTPERMAEAGFIHCPTENEPDLAQCFPC 60  
 QY 61 FKELEGWEPDDPIEBHKHSSGCAFLSVKKQFEBLTGFEFLKDRERAKNKIAKETNNK 120  
 DB 61 FKELEGWEPDDPIEBHKHSSGCAFLSVKKQFEBLTGFEFLKDRERAKNKIAKETNNK 120  
 QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
 DB 121 KKEFEETAKKVRRAIEQLAAMD 142  
 RESULT 9  
 AAU74573  
 ID AAU74573 standard; protein; 142 AA.  
 XX AAU74573;  
 AC  
 XX 08-MAY-2002 (first entry)  
 DT  
 XX Human mutant W10A survivin polypeptide.  
 DE  
 XX Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
 KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;  
 KW dimerisation; ovary; breast; pancreas; central nervous system; blood;  
 KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
 KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
 KW cell proliferative disorder; acquired immunodeficiency syndrome;  
 KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
 KW myelodysplastic syndrome; protein coordinate data; mutant; mutein.  
 KW

```

XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 10
XX FT /note= "Wild-type Trp substituted by Ala"
XX PN WO200202622-A2.
XX XX
XX PD 10-JAN-2002.
XX XX
XX PF 29-JUN-2001; 2001WO-US020872.
XX XX
XX PR 29-JUN-2000; 2000US-00608352.
XX XX
XX PA (SALK ) SALK INST BIOLOGICAL STUDIES.
XX PI Noel JP, Verdacia M, Hunter T, Huang H;
XX DR WPI; 2002-171638/22.
XX XX
XX PT Novel isolated crystalline survivin polypeptide, useful to screen and
XX PT design compounds that bind to or interact with inhibitor of apoptosis
XX PT protein and protein family members, and for design of novel therapeutics.
XX PS Claim 8; Page; 102pp; English.
XX XX
XX CC The invention relates to an isolated crystalline human survivin
XX CC polypeptide, an inhibitor or apoptosis protein (IAP). The polypeptide is
XX CC useful for determining the ability of a survivin binding agent to
XX CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc
XX CC chelation activity and for identifying an agent such as a peptide or
XX CC peptidomimetic which inhibits dimerisation of survivin. The ability of a
XX CC agent to modulate dimerisation can be determined through detection of a
XX CC change in apoptosis in a target cell expressing survivin. Binding agents
XX CC are useful for increasing apoptosis in a cell derived from a tissue
XX CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,
XX CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous
XX CC system, prostate, colon, rectum, cervix or endometrium, with a cell
XX CC proliferative disorder such as cancer. The molecules can also be used in
XX CC treatment or prevention of apoptosis which occurs as a part of AIDS,
XX CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease
XX CC and myelodysplastic syndromes. This sequence represents a human mutant
XX CC survivin polypeptide. Note: This sequence is not featured in the printed
XX CC specification but was derived from the wild type protein shown in
XX CC AA074571
XX SQ Sequence 142 AA;
XX
XX Query Match 93.0%; Score 132; DB 5; Length 142;
XX Best Local Similarity 100.0%; Pred. No. 4.6e-132; Indels 0; Gaps 0;
XX Matches 132; Conservative 0; Mismatches 0;
XX
XX QY 11 QPFLKDHRISTFKNWPFLGCGACTPERMAEAGFIHCPTNEPDLAQCFPCFKELEGWEPD 70
XX DB 11 QPFLKDHRISTFKNWPFLGCGACTPERMAEAGFIHCPTNEPDLAQCFPCFKELEGWEPD 70
XX
XX QY 71 DDFIEBKHSSGCAPLSVKKQFEELTLGFLKLDREAKNKIAKETNNKKKEPETAKK 130
XX DB 71 DDFIEBKHSSGCAPLSVKKQFEELTLGFLKLDREAKNKIAKETNNKKKEPETAKK 130
XX
XX QY 131 VRAIEQLAAMD 142
XX DB 131 VRAIEQLAAMD 142
XX
XX RESULT 10
XX ABUS6477
XX ID ABUS6477 standard; protein; 141 AA.
XX AC ABUS6477;
XX XX

```

```

DT 02-APR-2003 (first entry)
XX Lung cancer-associated polypeptide #70.
XX
XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
XX KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
XX KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
XX KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
XX KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
XX OS Unidentified.
XX PN WO200286443-A2.
XX PD 31-OCT-2002.
XX
XX PF 18-APR-2002; 2002WO-US012476.
XX PR 18-APR-2001; 2001US-0284770P.
XX PR 10-MAY-2001; 2001US-0290492P.
XX PR 09-NOV-2001; 2001US-0339245P.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 29-NOV-2001; 2001US-0334370P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX PI Aziz N, Murray R;
XX
XX DR WPI; 2003-093161/08.
XX DR N-PSDB; ABX76199.
XX
XX PT Detecting a lung cancer-associated transcript in a cell from a patient
XX PT for treating lung cancer, by contacting a biological sample from the
XX PT patient with a polynucleotide that exhibits increased or decreased
XX PT expression in lung cancer.
XX
XX PS Claim 27; Page 242; 453pp; English.
XX
XX CC The invention relates to a method for detecting a lung cancer-associated
XX CC transcript in a cell from a patient, comprising contacting a biological
XX CC sample from the patient with a polynucleotide that selectively hybridises
XX CC to a sequence that is at least 80 % identical to a gene that exhibits
XX CC increased or decreased expression in lung cancer samples. Lung cancer-
XX CC associated polynucleotides and polypeptides are used for identifying a
XX CC compound that modulates a lung cancer-associated polypeptide, for
XX CC inhibiting proliferation of a lung cancer-associated cell to treat lung
XX CC cancer in a patient and for treating a mammal having lung cancer by
XX CC administering a modulatory compound identified. The methods are useful
XX CC for treating lung cancer, such as small cell lung cancer, non-small cell
XX CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
XX CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
XX CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
XX CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
XX CC for diagnostic purposes and as targets for screening for therapeutic
XX CC compounds that modulate lung cancer, such as antibodies. Sequences
XX CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
XX CC invention
XX SQ Sequence 141 AA;
XX
XX Query Match 90.8%; Score 129; DB 6; Length 141;
XX Best Local Similarity 100.0%; Pred. No. 7e-129;
XX Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MGAPTLPPAWQPPFLKDHRISTFKNWPFLGCGACTPERMAEAGFIHCPTNEPDLAQCFPC 60
XX DB 1 MGAPTLPPAWQPPFLKDHRISTFKNWPFLGCGACTPERMAEAGFIHCPTNEPDLAQCFPC 60
XX
XX QY 61 FKELEGWEPDDPIEBKHSSGCAPLSVKKQFEELTLGFLKLDREAKNKIAKETNNK 120
XX DB 61 FKELEGWEPDDPIEBKHSSGCAPLSVKKQFEELTLGFLKLDREAKNKIAKETNNK 120

```



```

QY      121 KKEFEETAK 129
Db      121 KKEFEETAK 129

RESULT 11
ID      AAU74582 standard; protein; 125 AA.
XX
AC      AAU74582;
XX
DT      08-MAY-2002 (first entry)
XX
DE      Human deletion mutant delta126-142 survivin polypeptide.
XX
KW      Human; survivin; inhibitor of apoptosis protein; IAP; cytosstatic; p21;
KW      tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;
KW      dimerisation; ovary; breast; pancreas; central nervous system; blood;
KW      lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;
KW      thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;
KW      cell proliferative disorder; acquired immunodeficiency syndrome;
KW      neurodegenerative disease; ischaemic injury; toxin-induced liver disease;
KW      myelodysplastic syndrome; protein coordinate data; mutant; mutin.
XX
OS      Homo sapiens.
OS      Synthetic.
XX
PN      WO200202622-A2.
XX
PD      10-JAN-2002.
XX
PF      29-JUN-2001; 2001WO-US020872.
XX
PR      29-JUN-2000; 2000US-00608352.
XX
PA      (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
PI      Noel JP, Verdacia M, Hunter T, Huang H;
XX
PT      WPI; 2002-171638/22.
XX
DR      Novel isolated crystalline survivin polypeptide, useful to screen and
XX      design compounds that bind to or interact with inhibitor of apoptosis
XX      protein and protein family members, and for design of novel therapeutics.
XX
PS      Claim 8; Page; 102pp; English.
XX
CC      The invention relates to an isolated crystalline human survivin
XX      polypeptide, an inhibitor or apoptosis protein (IAP). The polypeptide is
XX      useful for determining the ability of a survivin binding agent to
XX      modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc
XX      chelation activity and for identifying an agent such as a peptide or
XX      peptidomimetic which inhibits dimerisation of survivin. The ability of an
XX      agent to modulate dimerisation can be determined through detection of a
XX      change in apoptosis in a target cell expressing survivin. Binding agents
XX      are useful for increasing apoptosis in a cell derived from a tissue
XX      selected from ovary, breast, pancreas, lymph node, skin, blood, lung,
XX      brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous
XX      system, prostate, colon, rectum, cervix or endometrium, with a cell
XX      treatment or prevention of apoptosis which occurs as a part of AIDS,
XX      neurodegenerative diseases, ischaemic injury, toxin-induced liver disease
XX      and myelodysplastic syndromes. This sequence represents a human deletion
XX      mutant survivin polypeptide, whereby residues 126-142 of the wild-type
XX      sequence have been deleted. Note: this sequence is not featured in the
XX      printed specification but was derived from the wild type protein shown in
XX      AAU74571
XX
SQ      Sequence 125 AA;
XX
Query Match      88.0%; Score 125; DB 5; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.1e-124; Indels 0; Gaps 0;
Matches 125; Conservative 0; Mismatches 0;

```

```

QY      1 MGAPTLPPAWQFFLKHHRISTFKWNPFLGCGACPTPERVABAGFIHCPTENPDLAQCFFC 60
Db      1 MGAPTLPPAWQFFLKHHRISTFKWNPFLGCGACPTPERVABAGFIHCPTENPDLAQCFFC 60

QY      61 FKELEGWEPDDDDPIEBHKHSSGCAFLSVKKQFBEFLTGLDRERAKNKIAKETNNK 120
Db      61 FKELEGWEPDDDDPIEBHKHSSGCAFLSVKKQFBEFLTGLDRERAKNKIAKETNNK 120

QY      121 KKEFE 125
Db      121 KKEFE 125

```

## RESULT 12

```

ID      AAU74574 standard; protein; 142 AA.
XX
AC      AAU74574;
XX
DT      08-MAY-2002 (first entry)
XX
DE      Human mutant T34E survivin polypeptide.
XX
KW      Human; survivin; inhibitor of apoptosis protein; IAP; cytosstatic; p21;
KW      tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;
KW      dimerisation; ovary; breast; pancreas; central nervous system; blood;
KW      lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;
KW      thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;
KW      cell proliferative disorder; acquired immunodeficiency syndrome;
KW      neurodegenerative disease; ischaemic injury; toxin-induced liver disease;
KW      myelodysplastic syndrome; protein coordinate data; mutant; mutin.
XX
OS      Homo sapiens.
OS      Synthetic.
XX
FH      Key Location/Qualifiers
FT      Misc-difference 34 /note= "Wild-type Thr substituted by Glu"
FT      WPI; 2002-171638/22.
XX
DN      WO200202622-A2.
XX
PD      10-JAN-2002.
XX
PF      29-JUN-2001; 2001WO-US020872.
XX
PR      29-JUN-2000; 2000US-00608352.
XX
PA      (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
PI      Noel JP, Verdacia M, Hunter T, Huang H;
XX
PT      WPI; 2002-171638/22.
XX
DR      Novel isolated crystalline survivin polypeptide, useful to screen and
XX      design compounds that bind to or interact with inhibitor of apoptosis
XX      protein and protein family members, and for design of novel therapeutics.
XX
PS      Claim 8; Page; 102pp; English.
XX
CC      The invention relates to an isolated crystalline human survivin
XX      polypeptide, an inhibitor or apoptosis protein (IAP). The polypeptide is
XX      useful for determining the ability of a survivin binding agent to
XX      modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc
XX      chelation activity and for identifying an agent such as a peptide or
XX      peptidomimetic which inhibits dimerisation of survivin. The ability of an
XX      agent to modulate dimerisation can be determined through detection of a
XX      change in apoptosis in a target cell expressing survivin. Binding agents
XX      are useful for increasing apoptosis in a cell derived from a tissue
XX      selected from ovary, breast, pancreas, lymph node, skin, blood, lung,
XX      brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous
XX      system, prostate, colon, rectum, cervix or endometrium, with a cell
XX      proliferative disorder such as cancer. The molecules can also be used in

```

CC treatment or prevention of apoptosis which occurs as a part of AIDS,  
 CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease  
 CC and myelodysplastic syndromes. This sequence represents a human mutant  
 CC survivin polypeptide. Note: This sequence is not featured in the printed  
 CC specification but was derived from the wild type protein shown in  
 CC AAU74571

XX SQ Sequence 142 AA;

Query Match 76.1%; Score 108; DB 5; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-106;  
 Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 35 PERMAEAGFTHCPTNEPDLAQCFKCELEGWEPDDPIEHHKSSGCAFLSVKKQFE 94  
 DB 35 PERMAEAGFTHCPTNEPDLAQCFKCELEGWEPDDPIEHHKSSGCAFLSVKKQFE 94  
 QY 95 ELTGLGEFLKDRERAKNIAKETNNKKKEFEETAKKVRRAIEQLAAMD 142  
 DB 95 ELTGLGEFLKDRERAKNIAKETNNKKKEFEETAKKVRRAIEQLAAMD 142

RESULT 13  
 AAU74577  
 ID AAU74577 standard; protein; 142 AA.

XX AC AAU74577;  
 XX DT 08-MAY-2002 (first entry)  
 XX DE Human mutant T97E survivin polypeptide.

XX KW Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
 KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;  
 KW dimerisation; ovary; breast; pancreas; central nervous system; blood;  
 KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
 KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
 KW cell proliferative disorder; acquired immunodeficiency syndrome;  
 KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
 KW myelodysplastic syndrome; protein coordinate data; mutant; mutein.

OS Homo sapiens.  
 OS Synthetic.

XX Key Location/Qualifiers

XX FT Misc-difference 97 /note= "Wild-type Thr substituted by Glu"

XX FN WO200202622-A2.

XX PD 10-JAN-2002.

XX PF 29-JUN-2001; 2001WO-US020872.

XX PR 29-JUN-2000; 2000US-00608352.

XX PA (SALK ) SALK INST BIOLOGICAL STUDIES.

XX PI Noel JP, Verdacia M, Hunter T, Huang H;

XX DR WPI; 2002-171638/22.

XX PT Novel isolated crystalline survivin polypeptide, useful to screen and  
 PT design compounds that bind to or interact with inhibitor of apoptosis  
 PT protein and protein family members, and for design of novel therapeutics.

XX PS Claim 8; Page; 102pp; English.

XX CC The invention relates to an isolated crystalline human survivin  
 CC polypeptide, an inhibitor or apoptosis protein (IAP). The polypeptide is  
 CC useful for determining the ability of a survivin binding agent to  
 CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc  
 CC chelation activity and for identifying an agent such as a peptide or

CC peptidomimetic which inhibits dimerisation of survivin. The ability of an  
 CC agent to modulate dimerisation can be determined through detection of a  
 CC change in apoptosis in a target cell expressing survivin. Binding agents  
 CC are useful for increasing apoptosis in a cell derived from a tissue  
 CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,  
 CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous  
 CC system, prostate, colon, rectum, cervix or endometrium, with a cell  
 CC proliferative disorder such as cancer. The molecules can also be used in  
 CC treatment or prevention of apoptosis which occurs as a part of AIDS,  
 CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease  
 CC and myelodysplastic syndromes. This sequence represents a human mutant  
 CC survivin polypeptide. Note: This sequence is not featured in the printed  
 CC specification but was derived from the wild type protein shown in  
 CC AAU74571

XX SQ Sequence 142 AA;

Query Match 67.6%; Score 96; DB 5; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 9.2e-94;  
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAPTLPPAWQPFLLKDHRISTFTKWPFLGCACTPERMAEAGFTHCPTNEPDLAQCFFC 60  
 DB 1 MGAPTLPPAWQPFLLKDHRISTFTKWPFLGCACTPERMAEAGFTHCPTNEPDLAQCFFC 60

QY 61 FKELEGWEPDDPIEHHKSSGCAFLSVKKQFEEL 96

DB 61 FKELEGWEPDDPIEHHKSSGCAFLSVKKQFEEL 96

RESULT 14

AAU74572  
 ID AAU74572 standard; protein; 142 AA.

XX AC AAU74572;

XX DT 08-MAY-2002 (first entry)

XX DE Human mutant L54M survivin polypeptide.

XX KW Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
 KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;  
 KW dimerisation; ovary; breast; pancreas; central nervous system; blood;  
 KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
 KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
 KW cell proliferative disorder; acquired immunodeficiency syndrome;  
 KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
 KW myelodysplastic syndrome; protein coordinate data; mutant; mutein.

OS Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

XX FT Misc-difference 54 /note= "Wild-type Leu substituted by Met"

XX FN WO200202622-A2.

XX PD 10-JAN-2002.

XX PF 29-JUN-2001; 2001WO-US020872.

XX PR 29-JUN-2000; 2000US-00608352.

XX PA (SALK ) SALK INST BIOLOGICAL STUDIES.

XX PI Noel JP, Verdacia M, Hunter T, Huang H;

XX DR WPI; 2002-171638/22.

XX PT Novel isolated crystalline survivin polypeptide, useful to screen and  
 PT design compounds that bind to or interact with inhibitor of apoptosis  
 PT protein and protein family members, and for design of novel therapeutics.

XX PS Claim 7; Page; 102pp; English.

XX CC The invention relates to an isolated crystalline human survivin

XX CC polypeptide, an inhibitor of apoptosis protein (IAP). The polypeptide is

XX CC useful for determining the ability of a survivin binding agent to

XX CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc

XX CC chelation activity and for identifying an agent such as a peptide or

XX CC peptidomimetic which inhibits dimerisation of survivin. The ability of an

XX CC agent to modulate dimerisation can be determined through detection of a

XX CC change in apoptosis in a target cell expressing survivin. Binding agents

XX CC are useful for increasing apoptosis in a cell derived from a tissue

XX CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,

XX CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous

XX CC system, prostate, colon, rectum, cervix or endometrium, with a cell

XX CC proliferative disorder such as cancer. The molecules can also be used in

XX CC treatment or prevention of apoptosis which occurs as a part of AIDS,

XX CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease

XX CC and myelodysplastic syndromes. This sequence represents a human mutant

XX CC survivin polypeptide. Note: This sequence is not featured in the printed

XX CC specification but was derived from the wild type protein shown in

XX CC AAU74571

XX SQ Sequence 142 AA;

Query Match 62.0%; Score 88; DB 5; Length 142;

Best Local Similarity 100.0%; Pred. No. 3e-85;

Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 55 AQCFFCPKELEGWEPDDPIEHHKXSSGCAFLSVKKQFEELTGLGFLKLDRAKNKIA 114

Db 55 AQCFFCPKELEGWEPDDPIEHHKXSSGCAFLSVKKQFEELTGLGFLKLDRAKNKIA 114

Qy 115 KETNNKKKEFEETAKVRRRAIEQLAAMD 142

Db 115 KETNNKKKEFEETAKVRRRAIEQLAAMD 142

RESULT 15

AAU74578

ID AAU74578 standard; protein; 142 AA.

AC AAU74578;

XX DT 08-MAY-2002 (first entry)

XX DE Human mutant W10A/L98A/F101R/L102S survivin polypeptide.

XX KW Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;

XX KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;

XX KW dimerisation; ovary; breast; pancreas; central nervous system; blood;

XX KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;

XX KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;

XX KW cell proliferative disorder; acquired immunodeficiency syndrome;

XX KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;

XX KW myelodysplastic syndrome; protein coordinate data; mutant; mutein.

XX OS Homo sapiens.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 10

FT /note= "Wild-type Trp substituted by Ala"

FT Misc-difference 98

FT /note= "Wild-type Leu substituted by Ala"

FT Misc-difference 101

FT /note= "Wild-type Phe substituted by Arg"

FT Misc-difference 102

FT /note= "Wild-type Leu substituted by Ser"

XX WO200202622-A2.

XX PN

XX 10-JAN-2002.

PD

XX PF 29-JUN-2001; 2001WO-US020872.

XX PR 29-JUN-2000; 2000US-00608352.

XX PA (SALK ) SALK INST BIOLOGICAL STUDIES.

XX PI Noel JP, Verdacia M, Hunter T, Huang H;

XX DR WPI; 2002-171638/22.

XX PT Novel isolated crystalline survivin polypeptide, useful to screen and

XX PT design compounds that bind to or interact with inhibitor of apoptosis

XX PT protein and protein family members, and for design of novel therapeutics.

XX PS Claim 8; Page; 102pp; English.

XX CC The invention relates to an isolated crystalline human survivin

XX CC polypeptide, an inhibitor of apoptosis protein (IAP). The polypeptide is

XX CC useful for determining the ability of a survivin binding agent to

XX CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc

XX CC chelation activity and for identifying an agent such as a peptide or

XX CC peptidomimetic which inhibits dimerisation of survivin. The ability of an

XX CC agent to modulate dimerisation can be determined through detection of a

XX CC change in apoptosis in a target cell expressing survivin. Binding agents

XX CC are useful for increasing apoptosis in a cell derived from a tissue

XX CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,

XX CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous

XX CC system, prostate, colon, rectum, cervix or endometrium, with a cell

XX CC proliferative disorder such as cancer. The molecules can also be used in

XX CC treatment or prevention of apoptosis which occurs as a part of AIDS,

XX CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease

XX CC and myelodysplastic syndromes. This sequence represents a human mutant

XX CC survivin polypeptide. Note: This sequence is not featured in the printed

XX CC specification but was derived from the wild type protein shown in

XX CC AAU74571

XX SQ Sequence 142 AA;

Query Match 61.3%; Score 87; DB 5; Length 142;

Best Local Similarity 100.0%; Pred. No. 3.5e-84;

Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 OPELKHRISTFKNPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFFCKELEGWEPD 70

Db 11 OPELKHRISTFKNPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFFCKELEGWEPD 70

Qy 71 DDPIEHHKXSSGCAFLSVKKQFEELT 97

Db 71 DDPIEHHKXSSGCAFLSVKKQFEELT 97

RESULT 16

AAU74579

ID AAU74579 standard; protein; 142 AA.

XX AC AAU74579;

XX DT 08-MAY-2002 (first entry)

XX DE Human mutant L6G/W10A/L98A/F101R/L102S survivin polypeptide.

XX KW Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;

XX KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;

XX KW dimerisation; ovary; breast; pancreas; central nervous system; blood;

XX KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;

XX KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;

XX KW cell proliferative disorder; acquired immunodeficiency syndrome;

XX KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;

XX KW myelodysplastic syndrome; protein coordinate data; mutant; mutein.

XX OS Homo sapiens.

XX OS Synthetic.

XX PH Key Location/Qualifiers  
 FT Misc-difference 6 /note= "Wild-type Leu substituted by Gly"  
 FT Misc-difference 10 /note= "Wild-type Trp substituted by Ala"  
 FT Misc-difference 98 /note= "Wild-type Leu substituted by Ala"  
 FT Misc-difference 101 /note= "Wild-type Phe substituted by Arg"  
 FT Misc-difference 102 /note= "Wild-type Leu substituted by Ser"  
 XX WO200202622-A2.  
 XX 10-JAN-2002.  
 XX 29-JUN-2001; 2001WO-US020872.  
 XX 29-JUN-2000; 2000US-00608352.  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX Noel JP, Verdacia M, Hunter T, Huang H;  
 XX WPI; 2002-171638/22.  
 XX Novel isolated crystalline survivin polypeptide, useful to screen and design compounds that bind to or interact with inhibitor of apoptosis protein and protein family members, and for design of novel therapeutics.  
 XX Claim 8; Page; 102pp; English.  
 XX The invention relates to an isolated crystalline human survivin polypeptide, an inhibitor of apoptosis protein (IAP). The polypeptide is useful for determining the ability of a survivin binding agent to modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc chelation activity and for identifying an agent such as a peptide or peptidomimetic which inhibits dimerisation of survivin. The ability of an agent to modulate dimerisation can be determined through detection of a change in apoptosis in a target cell expressing survivin. Binding agents are useful for increasing apoptosis in a cell derived from a tissue selected from ovary, breast, pancreas, lymph node, skin, blood, lung, brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous system, prostate, colon, rectum, cervix or endometrium, with a cell proliferative disorder such as cancer. The molecules can also be used in treatment or prevention of apoptosis which occurs as a part of AIDS, neurodegenerative diseases, ischaemic injury, toxin-induced liver disease and myelodysplastic syndromes. This sequence represents a human mutant survivin polypeptide. Note: This sequence is not featured in the printed specification but was derived from the wild type protein shown in AAU74571  
 XX Sequence 142 AA;  
 Query Match 61.3%; Score 87; DB 5; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-94;  
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 QPFLKDHRISTFKNWPFLGCACTPERMAEAGFHCTENEPDLAQCFKLEGWEPD 70  
 DB 11 QPFLKDHRISTFKNWPFLGCACTPERMAEAGFHCTENEPDLAQCFKLEGWEPD 70  
 QY 71 DPIEKKHSGGCAFLSVKKQFELT 97  
 DB 71 DPIEKKHSGGCAFLSVKKQFELT 97  
 RESULT 17  
 AAU74580  
 ID AAU74580 standard; protein; 142 AA.  
 XX  
 AC AAU74580;

XX DT 08-MAY-2002 (first entry)  
 XX DE Human mutant W10A/F93A/L98R survivin polypeptide.  
 XX KW Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21; tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis; dimerisation; ovary; breast; pancreas; central nervous system; blood; lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity; thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS; cell proliferative disorder; acquired immunodeficiency syndrome; neurodegenerative disease; ischaemic injury; toxin-induced liver disease; myelodysplastic syndrome; protein coordinate data; mutant; mutain.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PH Key Location/Qualifiers  
 FT Misc-difference 10 /note= "Wild-type Trp substituted by Ala"  
 FT Misc-difference 93 /note= "Wild-type Phe substituted by Ala"  
 FT Misc-difference 98 /note= "Wild-type Leu substituted by Arg"  
 XX WO200202622-A2.  
 XX 10-JAN-2002.  
 XX 29-JUN-2001; 2001WO-US020872.  
 XX 29-JUN-2000; 2000US-00608352.  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX Noel JP, Verdacia M, Hunter T, Huang H;  
 XX WPI; 2002-171638/22.  
 XX Novel isolated crystalline survivin polypeptide, useful to screen and design compounds that bind to or interact with inhibitor of apoptosis protein and protein family members, and for design of novel therapeutics.  
 XX Claim 8; Page; 102pp; English.  
 XX The invention relates to an isolated crystalline human survivin polypeptide, an inhibitor of apoptosis protein (IAP). The polypeptide is useful for determining the ability of a survivin binding agent to modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc chelation activity and for identifying an agent such as a peptide or peptidomimetic which inhibits dimerisation of survivin. The ability of an agent to modulate dimerisation can be determined through detection of a change in apoptosis in a target cell expressing survivin. Binding agents are useful for increasing apoptosis in a cell derived from a tissue selected from ovary, breast, pancreas, lymph node, skin, blood, lung, brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous system, prostate, colon, rectum, cervix or endometrium, with a cell proliferative disorder such as cancer. The molecules can also be used in treatment or prevention of apoptosis which occurs as a part of AIDS, neurodegenerative diseases, ischaemic injury, toxin-induced liver disease and myelodysplastic syndromes. This sequence represents a human mutant survivin polypeptide. Note: This sequence is not featured in the printed specification but was derived from the wild type protein shown in AAU74571  
 XX Sequence 142 AA;  
 Query Match 57.7%; Score 82; DB 5; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-79;  
 Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 QPFLKDHRISTFKNWPFLGCACTPERMAEAGFHCTENEPDLAQCFKLEGWEPD 70  
 DB 11 QPFLKDHRISTFKNWPFLGCACTPERMAEAGFHCTENEPDLAQCFKLEGWEPD 70  
 QY 71 DPIEKKHSGGCAFLSVKKQFELT 97  
 DB 71 DPIEKKHSGGCAFLSVKKQFELT 97  
 RESULT 17  
 AAU74580  
 ID AAU74580 standard; protein; 142 AA.  
 XX  
 AC AAU74580;

Db 11 QPFLKDHRISTFKNPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFCKELEGWEPD 70  
 QY 71 DDPIEHKHHSSGCAFLSVKKQ 92  
 Db 71 DDPIEHKHHSSGCAFLSVKKQ 92

RESULT 18  
 AAU74581  
 ID AAU74581 standard; protein; 142 AA.  
 XX AC AAU74581;  
 XX DT 08-MAY-2002 (first entry)  
 XX DE Human mutant L6G/W10A/F93A/L98R survivin polypeptide.  
 XX KW Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
 KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;  
 KW dimerisation; ovary; breast; pancreas; central nervous system; blood;  
 KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
 KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
 KW cell proliferative disorder; acquired immunodeficiency syndrome;  
 KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
 KW myelodysplastic syndrome; protein coordinate data; mutant; muten.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 6 /note= "Wild-type Leu substituted by Gly"  
 FT FT Misc-difference 10 /note= "Wild-type Trp substituted by Ala"  
 FT FT Misc-difference 93 /note= "Wild-type Phe substituted by Ala"  
 FT FT Misc-difference 98 /note= "Wild-type Leu substituted by Arg"  
 FT FN WO200202622-A2.  
 XX 10-JAN-2002.  
 XX 29-JUN-2001; 2001WO-US020872.  
 XX 29-JUN-2000; 2000US-00608352.  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX Noel JP, Verdacia M, Hunter T, Huang H;  
 XX WPI; 2002-171638/22.  
 XX Novel isolated crystalline survivin polypeptide, useful to screen and  
 PT design compounds that bind to or interact with inhibitor of apoptosis  
 PT protein and protein family members, and for design of novel therapeutics.  
 XX Claim 8; Page; 102pp; English.  
 XX The invention relates to an isolated crystalline human survivin  
 CC polypeptide, an inhibitor of apoptosis protein (IAP). The polypeptide is  
 CC useful for determining the ability of a survivin binding agent to  
 CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc  
 CC chelation activity and for identifying an agent such as a peptide or  
 CC peptidomimetic which inhibits dimerisation of survivin. The ability of an  
 CC agent to modulate dimerisation can be determined through detection of a  
 CC change in apoptosis in a target cell expressing survivin. Binding agents  
 CC are useful for increasing apoptosis in a cell derived from a tissue  
 CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,  
 CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous  
 CC system, prostate, colon, rectum, cervix or endometrium, with a cell  
 CC proliferative disorder such as cancer. The molecules can also be used in  
 CC treatment or prevention of apoptosis which occurs as a part of AIDS,

CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease  
 CC and myelodysplastic syndromes. This sequence represents a human mutant  
 CC survivin polypeptide. Note: This sequence is not featured in the printed  
 CC specification but was derived from the wild type protein shown in  
 CC AAU74571  
 XX SQ Sequence 142 AA;  
 Query Match 57.7%; Score 82; DB 5; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-79;  
 Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 QPFLKDHRISTFKNPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFCKELEGWEPD 70  
 Db 11 QPFLKDHRISTFKNPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFCKELEGWEPD 70  
 QY 71 DDPIEHKHHSSGCAFLSVKKQ 92  
 Db 71 DDPIEHKHHSSGCAFLSVKKQ 92

RESULT 19  
 AAU74575  
 ID AAU74575 standard; protein; 142 AA.  
 XX AC AAU74575;  
 XX DT 08-MAY-2002 (first entry)  
 XX DE Human mutant H80A survivin polypeptide.  
 XX KW Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
 KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;  
 KW dimerisation; ovary; breast; pancreas; central nervous system; blood;  
 KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
 KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
 KW cell proliferative disorder; acquired immunodeficiency syndrome;  
 KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
 KW myelodysplastic syndrome; protein coordinate data; mutant; muten.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 80 /note= "Wild-type His substituted by Ala"  
 FT FN WO200202622-A2.  
 XX 10-JAN-2002.  
 XX 29-JUN-2001; 2001WO-US020872.  
 XX 29-JUN-2000; 2000US-00608352.  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX Noel JP, Verdacia M, Hunter T, Huang H;  
 XX WPI; 2002-171638/22.  
 XX Novel isolated crystalline survivin polypeptide, useful to screen and  
 PT design compounds that bind to or interact with inhibitor of apoptosis  
 PT protein and protein family members, and for design of novel therapeutics.  
 XX Claim 8; Page; 102pp; English.  
 XX The invention relates to an isolated crystalline human survivin  
 CC polypeptide, an inhibitor of apoptosis protein (IAP). The polypeptide is  
 CC useful for determining the ability of a survivin binding agent to  
 CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc  
 CC chelation activity and for identifying an agent such as a peptide or  
 CC peptidomimetic which inhibits dimerisation of survivin. The ability of an  
 CC agent to modulate dimerisation can be determined through detection of a  
 CC change in apoptosis in a target cell expressing survivin. Binding agents  
 CC are useful for increasing apoptosis in a cell derived from a tissue  
 CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,  
 CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous  
 CC system, prostate, colon, rectum, cervix or endometrium, with a cell  
 CC proliferative disorder such as cancer. The molecules can also be used in  
 CC treatment or prevention of apoptosis which occurs as a part of AIDS,

CC agent to modulate dimerisation can be determined through detection of a  
 CC change in apoptosis in a target cell expressing survivin. Binding agents  
 CC are useful for increasing apoptosis in a cell derived from a tissue  
 CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,  
 CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous  
 CC system, prostate, colon, rectum, cervix or endometrium, with a cell  
 CC proliferative disorder such as cancer. The molecules can also be used in  
 CC treatment or prevention of apoptosis which occurs as a part of AIDS,  
 CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease  
 CC and myelodysplastic syndromes. This sequence represents a human mutant  
 CC survivin polypeptide. Note: This sequence is not featured in the printed  
 CC specification but was derived from the wild type protein shown in  
 CC AAU74571

XX  
 SQ Sequence 142 AA;  
 Query Match 55.6%; Score 79; DB 5; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-75;  
 Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGAPTLPPAWQPFLLKDHRISTFKWPFLEGGCACTPERMAEAGFHCPTNEPDLAQCF 60  
 DB 1 MGAPTLPPAWQPFLLKDHRISTFKWPFLEGGCACTPERMAEAGFHCPTNEPDLAQCF 60  
 QY 61 FKELEGWEPDDPIEEKK 79  
 DB 61 FKELEGWEPDDPIEEKK 79

RESULT 20  
 AAG02311  
 ID AAG02311 standard; protein; 116 AA.  
 AC AAG02311;  
 DT 06-OCT-2000 (first entry)  
 DE Human secreted protein, SEQ ID NO: 6392.  
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping.  
 OS Homo sapiens.  
 XX EP1033401-A2.  
 XX 06-SEP-2000.  
 XX 21-FEB-2000; 2000EP-00200610.  
 XX 26-FEB-1999; 99US-0122487P.  
 XX (GEST ) GENSET.  
 XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 DR WPI; 2000-500381/45.  
 DR N-PSDB; AAC02317.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 XX  
 XX Claim 13; SEQ ID NO 6392; 71pp + Sequence Listing; English.  
 CC The present sequence is a polypeptide encoded by one of a large number of  
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were  
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different  
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated  
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT  
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA  
 CC sequences derived from the 5' ends of mRNAs and even in those cases where  
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely

CC included, 5' ESTs are derived from mRNAs with intact 5' ends and can  
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs  
 CC are also used in diagnostic, forensic, gene therapy and chromosome  
 CC mapping procedures. They are used to obtain upstream regulatory sequences  
 CC and to design expression and secretion vectors  
 XX  
 SQ Sequence 116 AA;

Query Match 53.5%; Score 76; DB 3; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-72;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGAPTLPPAWQPFLLKDHRISTFKWPFLEGGCACTPERMAEAGFHCPTNEPDLAQCF 60  
 DB 1 MGAPTLPPAWQPFLLKDHRISTFKWPFLEGGCACTPERMAEAGFHCPTNEPDLAQCF 60  
 QY 61 FKELEGWEPDDPIEE 76  
 DB 61 FKELEGWEPDDPIEE 76

RESULT 21  
 ADB61824  
 ID ADB61824 standard; protein; 75 AA.  
 AC ADB61824;  
 DT 04-DEC-2003 (first entry)  
 DE Human inhibitor of apoptosis (IAP) protein Survivin BIR3 domain.  
 KW baculovirus inhibitor of apoptosis repeat domain; BIR domain;  
 KW apoptosis pathway; embryonic development; viral pathogenesis; cancer;  
 KW autoimmune disorder; neurodegenerative disease; apoptotic response;  
 KW systemic lupus erythematosus; multiple sclerosis; viral infection;  
 KW herpes virus; poxvirus; adenovirus; inhibitor of apoptosis; IAP; XIAP;  
 KW HIAP1; CIAP2; HIAP2; CIAP1; RING zinc finger; caspase-3; caspase-7;  
 KW caspase-9; cytostatic; neoplasm; leukaemia; colon carcinoma;  
 KW cervical cancer; uterine cancer; testicular cancer;  
 KW small cell lung carcinoma; uterine cancer; renal cell carcinoma;  
 KW Wilm's tumour; human; BIR 3 domain.  
 XX Homo sapiens.  
 XX WO2003040172-A2.  
 XX 15-MAY-2003.  
 XX 12-NOV-2002; 2002WO-CA001738.  
 XX 09-NOV-2001; 2001US-032300P.  
 XX 08-APR-2002; 2002US-0370934P.  
 XX (AEGE-) AEGERA THERAPEUTICS INC.  
 XX Boudreaux A, Korneluk RG, La Casse E, Liston P;  
 XX WPI; 2003-513532/48.

XX Polypeptide capable of forming a complex with a polypeptide comprising a  
 PT baculovirus inhibitor of apoptosis repeat domain useful for treating  
 PT cancer and other neoplasms.  
 XX Disclosure; Fig 1A; 53pp; English.  
 CC This invention relates to a substantially pure polypeptide having a  
 CC length of less than 100 amino acids and capable of forming a complex with  
 CC a polypeptide that includes a baculovirus inhibitor of apoptosis repeat  
 CC (BIR) domain. The apoptosis pathway is known to play a critical role in  
 CC embryonic development, viral pathogenesis, cancer, autoimmune disorders  
 CC and neurodegenerative diseases. The failure of the apoptotic response has  
 CC been implicated in the development of cancer, autoimmune disorders (for  
 CC example systemic lupus erythematosus and multiple sclerosis) and viral

CC infections (including herpes virus, poxvirus and adenovirus. The

CC inhibitors of apoptosis (IAPs) are a family of proteins possessing one or

CC more baculovirus IAP repeat (BIR) domains. Human IAPs, XIAP, HIAP1

CC (CIAP2) and HIAP2 (CIAP1) all possess three BIR domains and carboxy

CC terminal RING zinc fingers. The IAPs bind and inhibit caspases -3, -7 and

CC -9 which are proteases involved in the initiation of apoptosis. Compounds

CC which inhibit the activity of IAPs may therefore have cytostatic activity

CC through the enhancement of apoptosis. The polypeptides of the invention

CC are candidate peptide ligands for binding to the BIR domain of IAPs. They

CC may be useful for the treatment of cancer and other neoplasms, such as

CC leukaemias, colon carcinoma, cervical cancer, uterine cancer, testicular

CC cancer, small cell lung carcinoma, uterine cancer, renal cell carcinoma

CC and Wilms' tumour, and for enhancing apoptosis. The present sequence is

CC that of the human inhibitor of apoptosis (IAP) protein Survivin BIR3

CC domain which was used to demonstrate homology to other human IAP protein

CC BIR domains to which the peptides of the invention are targeted to bind.

XX

XX Sequence 75 AA;

Query Match 52.8%; Score 75; DB 7; Length 75;

Best Local Similarity 100.0%; Pred. No. 1.1e-71;

Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 KDHRISTFKNWPFLGCACTPERMAEAGFIHCTPTENEDPLAQCFFCFKLEGWEPDDDDPI 74

DB 1 KDHRISTFKNWPFLGCACTPERMAEAGFIHCTPTENEDPLAQCFFCFKLEGWEPDDDDPI 60

QY 75 EEHKHSSGCAFLSV 89

DB 61 EEHKHSSGCAFLSV 75

RESULT 22

AAU74576

ID AAU74576 standard; protein; 142 AA.

XX AAU74576;

XX 08-MAY-2002 (first entry)

XX Human mutant H80A/E76A survivin polypeptide.

XX Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;

XX tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;

XX dimerisation; ovary; breast; pancreas; central nervous system; blood;

XX lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;

XX thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;

XX cell proliferative disorder; acquired immunodeficiency syndrome;

XX neurodegenerative disease; ischaemic injury; toxin-induced liver disease;

XX myelodysplastic syndrome; protein coordinate data; mutant; mutein.

OS Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

XX Misc-difference 76

XX /note= "Wild-type Glu substituted by Ala"

XX Misc-difference 80

XX /note= "Wild-type His substituted by Ala"

XX WO200202622-A2.

XX 10-JAN-2002.

XX 29-JUN-2001; 2001WO-US020872.

XX 29-JUN-2000; 2000US-00608352.

XX (SALK ) SALK INST BIOLOGICAL STUDIES.

XX Noel JP, Verdacia M, Hunter T, Huang H;

XX WPI; 2002-171638/22.

DR

XX Novel isolated crystalline survivin polypeptide, useful to screen and

PT design compounds that bind to or interact with inhibitor of apoptosis

PT protein and protein family members, and for design of novel therapeutics.

XX

PS Claim 8; Page; 102pp; English.

XX The invention relates to an isolated crystalline human survivin

CC polypeptide, an inhibitor or apoptosis protein (IAP). The polypeptide is

CC useful for determining the ability of a survivin binding agent to

CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc

CC chelation activity and for identifying an agent such as a peptide or

CC peptidomimetic which inhibits dimerisation of survivin. The ability of an

CC agent to modulate dimerisation can be determined through detection of a

CC change in apoptosis in a target cell expressing survivin. Binding agents

CC are useful for increasing apoptosis in a cell derived from a tissue

CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,

CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous

CC system, prostate, colon, rectum, cervix or endometrium, with a cell

CC proliferative disorder such as cancer. The molecules can also be used in

CC treatment or prevention of apoptosis which occurs as a part of AIDS,

CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease

CC and myelodysplastic syndromes. This sequence represents a human mutant

CC survivin polypeptide. Note: This sequence is not featured in the printed

CC specification but was derived from the wild type protein shown in

CC AAU74571

XX

XX Sequence 142 AA;

Query Match 52.8%; Score 75; DB 5; Length 142;

Best Local Similarity 100.0%; Pred. No. 2e-71;

Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAPTLPPAWQPLFKDHRISTFKNWPFLGCACTPERMAEAGFIHCTPTENEDPLAQCFFC 60

DB 1 MGAPTLPPAWQPLFKDHRISTFKNWPFLGCACTPERMAEAGFIHCTPTENEDPLAQCFFC 60

QY 61 FKELEGWEPDDDDPIE 75

DB 61 FKELEGWEPDDDDPIE 75

RESULT 23

ADA09936

ID ADA09936 standard; protein; 165 AA.

XX ADA09936;

XX 06-NOV-2003 (first entry)

XX Human survivin, SEQ ID NO:4.

XX Human; survivin; survivin 2B80; HLA-A24 antigen; human leukocyte antigen;

XX cytotoxic T-cell; CTL; induction; cancer; vaccine; anticancer agent;

XX cytostatic.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 80..88

XX /label= Survivin\_2B80

XX /note= "Corresponds to SEQ ID NO:1, which is specifically

XX claimed"

XX Peptide 108..116

XX /label= Survivin\_85

XX /note= "Corresponds to SEQ ID NO:2"

XX Peptide 115..124

XX /label= Survivin\_92

XX /note= "Corresponds to SEQ ID NO:3"

XX

PN JP2002284797-A.

XX

XX 03-OCT-2002.

XX PF 23-MAR-2001; 2001JP-00084438.  
 XX PF 23-MAR-2001; 2001JP-00084438.  
 XX PA (HOKK-) HOKKAIDO TLO KK.  
 XX DR WPI; 2003-397157/38.  
 XX PT A novel nonapeptide survivin 2B80 capable of inducing cytotoxic T-cells,  
 XX PT useful in vaccines for the treatment of cancer.  
 XX PS Example 3; Page 4-5; 10pp; Japanese.  
 XX CC The invention relates to a human survivin-derived cancer antigen,  
 XX CC designated survivin 2B80. Survivin 2B80 is a HLA-A24 antigen, and is  
 XX CC capable of inducing cytotoxic T-cells (CTLs) which target cancer cells  
 XX CC having HLA-A24 expression. Survivin 2B80 can be used in the treatment of  
 XX CC cancer, specifically in a cancer vaccine or as an anticancer agent. The  
 XX CC present sequence represents full-length human survivin which was used in  
 XX CC an example from the invention.  
 XX SQ Sequence 165 AA;  
 Query Match 52.1%; Score 74; DB 6; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-70;  
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGAPTLPPAWOPFLKDRHSTFKWPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCF 60  
 DB 1 MGAPTLPPAWOPFLKDRHSTFKWPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCF 60  
 QY 61 FKELEGWEPDDPI 74  
 DB 61 FKELEGWEPDDPI 74  
 RESULT 24  
 ADB61823  
 ID ADB61823 standard; protein; 83 AA.  
 AC ADB61823;  
 DT 04-DEC-2003 (first entry)  
 DE Human inhibitor of apoptosis (IAP) protein BC000784 BIR3 domain.  
 KW baculovirus inhibitor of apoptosis repeat domain; BIR domain;  
 KW apoptosis pathway; embryonic development; viral pathogenesis; cancer;  
 KW autoimmune disorder; neurodegenerative disease; apoptotic response;  
 KW systemic lupus erythematosus; multiple sclerosis; viral infection;  
 KW herpes virus; poxvirus; adenovirus; inhibitor of apoptosis; IAP; XIAP;  
 KW XIAP2; XIAP2; CIAP1; RING zinc finger; caspase-3; caspase-7;  
 KW caspase-9; cytoskeletal; neoplasm; leukaemia; colon carcinoma;  
 KW cervical cancer; uterine cancer; testicular cancer;  
 KW small cell lung carcinoma; uterine cancer; renal cell carcinoma;  
 KW Wilm's tumour; human; BIR 3 domain.  
 XX Homo sapiens.  
 XX WO2003040172-A2.  
 XX 15-MAY-2003.  
 XX 12-NOV-2002; 2002WO-CA001738.  
 XX 09-NOV-2001; 2001US-0322300P.  
 XX 08-APR-2002; 2002US-0370934P.  
 XX (AEGE-) AEGERA THERAPEUTICS INC.  
 XX Boudreault A, Korneluk RG, La Casse E, Liston P;  
 XX

DR WPI; 2003-513532/48.  
 XX Polypeptide capable of forming a complex with a polypeptide comprising a  
 XX PT baculovirus inhibitor of apoptosis repeat domain useful for treating  
 XX PT cancer and other neoplasms.  
 XX PS Disclosure; Fig 1A; 53pp; English.  
 XX CC This invention relates to a substantially pure polypeptide having a  
 XX CC length of less than 100 amino acids and capable of forming a complex with  
 XX CC a polypeptide that includes a baculovirus inhibitor of apoptosis repeat  
 XX CC (BIR) domain. The apoptosis pathway is known to play a critical role in  
 XX CC embryonic development, viral pathogenesis, cancer, autoimmune disorders  
 XX CC and neurodegenerative diseases. The failure of the apoptotic response has  
 XX CC been implicated in the development of cancer, autoimmune disorders (for  
 XX CC example systemic lupus erythematosus and multiple sclerosis) and viral  
 XX CC infections (including herpes virus, poxvirus and adenovirus). The  
 XX CC inhibitors of apoptosis (IAPs) are a family of proteins possessing one or  
 XX CC more baculovirus IAP repeat (BIR) domains. Human IAPs, XIAP, HIAP1  
 XX CC (CIAP2) and HIAP2 (CIAP1) all possess three BIR domains and carboxy  
 XX CC terminal RING zinc fingers. The IAPs bind and inhibit caspases -3, -7 and  
 XX CC -9 which are proteases involved in the initiation of apoptosis. Compounds  
 XX CC which inhibit the activity of IAPs may therefore have cytostatic activity  
 XX CC through the enhancement of apoptosis. The polypeptides of the invention  
 XX CC are candidate peptide ligands for binding to the BIR domain of IAPs. They  
 XX CC may be useful for the treatment of cancer and other neoplasms, such as  
 XX CC leukaemias, colon carcinoma, cervical cancer, uterine cancer, testicular  
 XX CC cancer, small cell lung carcinoma, uterine cancer, renal cell carcinoma  
 XX CC and Wilm's tumour, and for enhancing apoptosis. The present sequence is  
 XX CC that of the human inhibitor of apoptosis (IAP) protein BC000784 BIR3  
 XX CC domain which was used to demonstrate homology to other human IAP protein  
 XX CC BIR domains to which the peptides of the invention are targeted to bind.  
 XX SQ Sequence 83 AA;  
 Query Match 41.5%; Score 59; DB 7; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-54;  
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 15 KDHRISTFKWPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFCKFKELEGWEPDDDP 73  
 DB 1 KDHRISTFKWPFLEGGCACTPERMAEAGFIHCPTENEPDLAQCFCKFKELEGWEPDDDP 59  
 RESULT 25  
 AAW19749  
 ID AAW19749 standard; protein; 140 AA.  
 AC AAW19749;  
 DT 16-SEP-1997 (first entry)  
 DE Mouse inhibitor of apoptosis protein homologue MIHE.  
 KW Inhibitor of apoptosis protein; IAP; mammalian IAP homologue; MIHE;  
 KW degenerative disease; infectious disease; autoimmune disease; cancer;  
 KW therapy; diagnosis.  
 XX Mus musculus.  
 XX WO9723501-A1.  
 XX 03-JUL-1997.  
 XX 20-DEC-1996; 96WO-AU000827.  
 XX 22-DEC-1995; 95AU-00007275.  
 XX (AMRA-) AMRAD OPERATIONS PTY LTD.  
 XX Vaux DL;  
 XX WPI; 1997-350966/32.  
 DR



DR N-PSDB; AAT72714.  
 XX Isolated protein homologues of viral inhibitors of apoptosis - used to  
 PT modulate apoptosis for treatment of degenerative, infectious or  
 PT auto-immune diseases and cancer.  
 XX  
 XX  
 PS Claim 11; Page 71-72; 136pp; English.  
 XX  
 CC MIHE (AAW19749) is a murine homologue of baculovirus inhibitor of  
 CC apoptosis protein (IAP). Its amino acid sequence was deduced from an  
 CC isolated nucleic acid (see also AAT72714) obtd. by a database search for  
 CC sequences homologous to a baculovirus IAP repeat (BIR) consensus sequence  
 CC (see also AAW19744). Unlike IAP, MIHE does not contain a RING finger  
 CC domain. IAP homologues (see also AAW19745-48 and AAW19750-52) and their  
 CC derivatives and chemical analogues can be used in methods for modulating  
 CC apoptosis in animal cells, specifically for treatment, by inhibition, of  
 CC degenerative and infectious disease or, by promotion, of cancer and  
 CC autoimmune disease  
 XX  
 SQ Sequence 140 AA;  
 Query Match 28.9%; Score 41; DB 2; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 3e-35;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 31 CACTPERMAEAGFIHCPTENEPLAQCFFCFKELEGWEPDD 71  
 DB 31 CACTPERMAEAGFIHCPTENEPLAQCFFCFKELEGWEPDD 71  
 RESULT 26  
 AAU74583  
 ID AAU74583 standard; protein; 140 AA.  
 AC AAU74583;  
 XX  
 XX  
 DT 08-MAY-2002 (first entry)  
 DE Mouse survivin polypeptide.  
 XX  
 XX Human; survivin; inhibitor of apoptosis protein; IAP; cytostatic; p21;  
 KW tubulin interaction; caspase-3; CDK4; zinc chelation activity; apoptosis;  
 KW dimersation; ovary; pancreas; central nervous system; blood;  
 KW lung; skin; lymph node; brain; kidney; liver; nasopharyngeal cavity;  
 KW thyroid; prostate; colon; rectum; cervix; endometrium; cancer; AIDS;  
 KW cell proliferative disorder; acquired immunodeficiency syndrome; mouse;  
 KW neurodegenerative disease; ischaemic injury; toxin-induced liver disease;  
 KW myelodysplastic syndrome; protein coordinate data.  
 XX  
 OS Mus sp.  
 XX  
 XX WO200202622-A2.  
 XX  
 PD 10-JAN-2002.  
 XX  
 XX  
 XX 29-JUN-2001; 2001WO-US020872.  
 XX  
 XX 29-JUN-2000; 2000US-00608352.  
 XX  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX  
 XX Noel JP, Verdacia M, Hunter T, Huang H;  
 XX WPI; 2002-171638/22.  
 XX  
 XX Novel isolated crystalline survivin polypeptide, useful to screen and  
 PT design compounds that bind to or interact with inhibitor of apoptosis  
 PT protein and protein family members, and for design of novel therapeutics.  
 XX  
 XX Disclosure; fig 2; 102pp; English.  
 PS  
 XX The invention relates to an isolated crystalline human survivin  
 CC polypeptide, an inhibitor or apoptosis protein (IAP). The polypeptide is

CC useful for determining the ability of a survivin binding agent to  
 CC modulate tubulin interaction, p21, caspase-3 and CDK4 requirement or zinc  
 CC chelation activity and for identifying an agent such as a peptide or  
 CC peptidomimetic which inhibits dimerisation of survivin. The ability of an  
 CC agent to modulate dimerisation can be determined through detection of a  
 CC change in apoptosis in a target cell expressing survivin. Binding agents  
 CC are useful for increasing apoptosis in a cell derived from a tissue  
 CC selected from ovary, breast, pancreas, lymph node, skin, blood, lung,  
 CC brain, kidney, liver, nasopharyngeal cavity, thyroid, central nervous  
 CC system, prostate, colon, rectum, cervix or endometrium, with a cell  
 CC proliferative disorder such as cancer. The molecules can also be used in  
 CC treatment or prevention of apoptosis which occurs as a part of AIDS,  
 CC neurodegenerative diseases, ischaemic injury, toxin-induced liver disease  
 CC and myelodysplastic syndromes. This sequence represents the mouse  
 CC survivin polypeptide  
 XX  
 SQ Sequence 140 AA;  
 Query Match 28.9%; Score 41; DB 5; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 3e-35;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 31 CACTPERMAEAGFIHCPTENEPLAQCFFCFKELEGWEPDD 71  
 DB 31 CACTPERMAEAGFIHCPTENEPLAQCFFCFKELEGWEPDD 71  
 RESULT 27  
 AAU30231  
 ID AAU30231 standard; protein; 155 AA.  
 AC AAU30231;  
 XX  
 XX 18-DEC-2001 (first entry)  
 DT  
 DE Novel human secreted protein #722.  
 XX  
 KW Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200179449-A2.  
 XX  
 PD 25-OCT-2001.  
 XX  
 XX 16-APR-2001; 2001WO-US008656.  
 XX  
 XX 18-APR-2000; 2000US-00552929.  
 XX  
 XX 26-JAN-2001; 2001US-00770160.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2001-611725/70.  
 XX  
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.  
 XX  
 XX Claim 20; Page 263; 765pp; English.  
 XX  
 XX The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins

CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX  
 SQ Sequence 155 AA;

Query Match 25.4%; Score 36; DB 4; Length 155;  
 Best Local Similarity 100.0%; Pred. No. 7e-30;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 MAEAGFIHCPTNEPDLAQCFCKELSGWEPDDDP 73  
 DB 1 MAEAGFIHCPTNEPDLAQCFCKELSGWEPDDDP 36

RESULT 28  
 ABP70506  
 ID ABP70506 standard; peptide; 24 AA.

XX AC ABP70506;  
 XX  
 DT 22-APR-2003 (first entry)  
 XX  
 DE Peptide comprising translocating motif, derived from survivin.  
 XX  
 KW Translocating factor; translocation; cell membrane; disease; tumour;  
 KW infection.  
 XX  
 OS Synthetic.

XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "biotin attached"  
 XX

XX PN WO2003002598-A2.  
 XX  
 XX PD 09-JAN-2003.

XX PF 01-JUL-2002; 2002WO-GB003027.

XX PR 29-JUN-2001; 2001GB-00016047.

XX PA (IMPL-) IMPLX LTD.

XX PI Crisanti A;

XX DR WPI; 2003-201490/19.

XX Use of a protein or its fragment for the manufacture of a composition for  
 PT treating a disease characterized by a deficiency in the production or  
 PT function of the endogenous protein, or to regulate biochemical pathway.

XX Example 2; Page 30; 44pp; English.

XX The present sequence represents a peptide comprising a translocating  
 CC factor motif. The peptide motif enables translocation across the cell  
 CC membrane. The conjugate was tested for translocation, which was observed.  
 CC The peptide motif is derived from a formula given in the specification.  
 CC The peptide motifs and proteins and conjugates comprising that motif are  
 CC useful for the manufacture of a medicament for treating or diagnosing a  
 CC disease, e.g. tumour, or viral and bacterial infections. The conjugate is  
 CC useful in therapy or diagnosis

XX Sequence 24 AA;

Query Match 16.9%; Score 24; DB 6; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 7.2e-18;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 111 NKIAKETNNKKKEFEETAKKVRRA 134  
 DB 1 NKIAKETNNKKKEFEETAKKVRRA 24

RESULT 29

ABP70505  
 ID ABP70505 standard; peptide; 21 AA.

XX AC ABP70505;

XX DT 22-APR-2003 (first entry)

XX DE Peptide comprising translocating motif, derived from baculoviral protein.  
 XX KW Translocating factor; translocation; cell membrane; disease; tumour;  
 XX KW infection.

XX OS Synthetic.

XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "biotin attached"

XX PN WO2003002598-A2.

XX PD 09-JAN-2003.

XX PF 01-JUL-2002; 2002WO-GB003027.

XX PR 29-JUN-2001; 2001GB-00016047.

XX PA (IMPL-) IMPLX LTD.

XX PI Crisanti A;

XX DR WPI; 2003-201490/19.

XX Use of a protein or its fragment for the manufacture of a composition for  
 PT treating a disease characterized by a deficiency in the production or  
 PT function of the endogenous protein, or to regulate biochemical pathway.

XX Example 2; Page 30; 44pp; English.

XX The present sequence represents a peptide comprising a translocating  
 CC factor motif. The peptide motif enables translocation across the cell  
 CC membrane. The conjugate was tested for translocation, which was observed.  
 CC The peptide motif is derived from a formula given in the specification.  
 CC The peptide motifs and proteins and conjugates comprising that motif are  
 CC useful for the manufacture of a medicament for treating or diagnosing a  
 CC disease, e.g. tumour, or viral and bacterial infections. The conjugate is  
 CC useful in therapy or diagnosis

XX Sequence 21 AA;

Query Match 14.8%; Score 21; DB 6; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 9.9e-15;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 AKETNNKKKEFEETAKKVRRA 134  
 DB 1 AKETNNKKKEFEETAKKVRRA 21

RESULT 30

AAW61358  
 ID AAW61358 standard; peptide; 20 AA.

XX AC AAW61358;

XX DT 25-SEP-1998 (first entry)

XX DE Survivin functionally relevant peptide.

XX survivin; apoptosis; cellular apoptosis; transplantation;  
 KW motor neuron degenerative disease; HIV infection; immunosuppression;  
 KW gastrointestinal perturbations; cardiovascular disorder.  
 XX Homo sapiens.  
 XX WO9822589-A2.  
 XX 28-MAY-1998.  
 XX 20-NOV-1997; 97WO-US021880.  
 XX 20-NOV-1996; 96US-0031435P.  
 XX 20-NOV-1997; 97US-00975080.  
 XX (UYVA ) UNIV YALE.  
 XX Altieri DC;  
 XX WPI; 1998-312475/27.  
 XX Modulating apoptosis by controlling the Survivin gene - useful for  
 PT treating transplant rejection, degenerative disorders and tumours.  
 XX Claim 7; Page 75; 108pp; English.  
 XX The survivin peptide is a functionally relevant area of the protein. A  
 CC mutation resulting in the substitutions of amino acids at sites 3, 9 and  
 CC 20 with Alanine results in a complete loss of function of survivin in  
 CC transfected cells. The survivin gene can be used to control apoptosis  
 CC through modification of the gene. Survivin peptides can be used to  
 CC inhibit cellular apoptosis, e.g. for enhancing the viability of organs  
 CC and tissues prior to their transplantation, for preserving the growth of  
 CC cells in culture or for treating conditions involving abnormal apoptosis,  
 CC e.g. degenerative diseases such as motor neuron degenerative diseases,  
 CC HIV infection, dermatological effects of ageing, disorders and diseases,  
 CC such as immunosuppression, gastrointestinal perturbations, cardiovascular  
 CC disorders, apoptosis related to reperfusion damage, rejection of tissue  
 CC transplantation and Alzheimer's disease. Agents which block Survivin  
 CC activity can be used to treat e.g. tumours  
 XX Sequence 20 AA;  
 Query Match 14.1%; Score 20; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 65 EGVPEDDPPIEHHKHSSGC 84  
 Db 1 EGVPEDDPPIEHHKHSSGC 20  
 RESULT 31  
 AA61360  
 ID AA61360 standard; peptide; 17 AA.  
 XX AC AA61360;  
 XX 25-SEP-1998 (first entry)  
 XX Homo sapiens.  
 XX Survivin seventeenmer peptide.  
 KW survivin; apoptosis; cellular apoptosis; transplantation;  
 KW motor neuron degenerative disease; HIV infection; immunosuppression;  
 KW gastrointestinal perturbations; cardiovascular disorder.  
 XX Homo sapiens.  
 XX WO9822589-A2.  
 XX 28-MAY-1998.  
 XX (UYVA ) UNIV YALE.

PF 20-NOV-1997; 97WO-US021880.  
 XX 20-NOV-1996; 96US-0031435P.  
 XX 20-NOV-1997; 97US-00975080.  
 XX (UYVA ) UNIV YALE.  
 XX Altieri DC;  
 XX WPI; 1998-312475/27.  
 XX Modulating apoptosis by controlling the Survivin gene - useful for  
 PT treating transplant rejection, degenerative disorders and tumours.  
 XX Example 2; Page 51; 108pp; English.  
 XX The seventeenmer peptide was used to produce a survivin sequence-specific  
 CC antibody. The survivin gene can be used to control apoptosis through  
 CC modification of the gene. Survivin peptides can be used to inhibit  
 CC cellular apoptosis, e.g. for enhancing the viability of organs and  
 CC tissues prior to their transplantation, for preserving the growth of  
 CC cells in culture or for treating conditions involving abnormal apoptosis,  
 CC e.g. degenerative diseases such as motor neuron degenerative diseases,  
 CC HIV infection, dermatological effects of ageing, disorders and diseases,  
 CC such as immunosuppression, gastrointestinal perturbations, cardiovascular  
 CC disorders, apoptosis related to reperfusion damage, rejection of tissue  
 CC transplantation and Alzheimer's disease. Agents which block Survivin  
 CC activity can be used to treat e.g. tumours  
 XX Sequence 17 AA;  
 Query Match 12.0%; Score 17; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 APTLPPAWQPFLLKHRI 19  
 Db 1 APTLPPAWQPFLLKHRI 17  
 RESULT 32  
 AA611066  
 ID AA611066 standard; peptide; 17 AA.  
 XX AC AA611066;  
 XX 18-DEC-2001 (first entry)  
 XX Human survivin peptide.  
 KW Survivin dependent apoptosis; p34(cdc2)-cyclin B1 kinase complex;  
 KW caspase-9; tumour suppressor; antisense therapy; cytotatic;  
 KW human. inhibitor of apoptosis; IAP.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Modified-site 7.10  
 FT /note= "Phosphorylation site for p34(cdc2)-cyclin B1  
 FT kinase complex"  
 FT Modified-site 7  
 FT /note= "Phosphorylated Thr"  
 XX WO200164741-A2.  
 XX 07-SEP-2001.  
 XX 28-FEB-2001; 2001WO-US006341.  
 XX 29-FEB-2000; 2000US-00515514.  
 XX (UYVA ) UNIV YALE.

PI Altieri DC;  
 XX WPI; 2001-589861/66.  
 XX Modulating survivin apoptosis pathways by using an agent that modulates  
 PT phosphorylation of survivin, interaction between survivin and p34-cyclin  
 PT B1 kinase complex and interaction between survivin and caspase-9.  
 XX  
 PS Claim 7; Page 45; 97pp; English.  
 XX  
 CC The invention relates to a method for modulating apoptosis, in particular  
 CC survivin dependent apoptosis in a cell. The method involves administering  
 CC to the cell an agent which modulates the interaction between survivin and  
 CC p34(cdc2)-cyclin B1 kinase complex, phosphorylation of survivin by the  
 CC p34(cdc2)-cyclin B1 kinase complex and interaction between survivin and  
 CC caspase-9. Survivin antagonist is useful for inhibiting tumour growth in  
 CC a mammal. Agents that modulate, preferably dephosphorylate survivin at  
 CC Thr34, act as potential tumour suppressors and prevent the formation of  
 CC anti-apoptotic complex with caspase. The method is also useful in  
 CC antisense therapy. The present sequence is human survivin peptide. This  
 CC peptide functions as a novel inhibitor of apoptosis (IAP)  
 XX  
 XX Sequence 17 AA;  
 SQ  
 Query Match 12.0%; Score 17; DB 4; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 28 LEGACTPERMAGRI 44  
 DB 1 LEGACTPERMAGRI 17  
 RESULT 33  
 ADA09938  
 ID ADA09938 standard; peptide; 10 AA.  
 AC ADA09938;  
 XX  
 XX 06-NOV-2003 (first entry)  
 DT Human HLA-A24 antigen survivin 92, SEQ ID NO:3.  
 DE  
 XX Human; survivin; survivin 2880; HLA-A24 antigen;  
 XX human leukocyte antigen; cytotoxic T-cell; CTL; induction; cancer;  
 XX vaccine; anticancer agent; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 XX JP2002284797-A.  
 XX  
 XX 03-OCT-2002.  
 XX  
 XX 23-MAR-2001; 2001JP-00084438.  
 XX  
 XX 23-MAR-2001; 2001JP-00084438.  
 XX  
 XX (HOKK-) HOKKAIDO TLO KK.  
 XX  
 XX WPI; 2003-397157/38.  
 XX  
 XX A novel nonapeptide survivin 2880 capable of inducing cytotoxic T-cells,  
 PT useful in vaccines for the treatment of cancer.  
 XX  
 PS Example 3; Page 5; 10pp; Japanese.  
 XX  
 CC The invention relates to a human survivin-derived cancer antigen,  
 CC designated survivin 2880. Survivin 2880 is a HLA-A24 antigen, and is  
 CC capable of inducing cytotoxic T-cells (CTLs) which target cancer cells  
 CC having HLA-A24 expression. Survivin 2880 can be used in the treatment of  
 CC cancer, specifically in a cancer vaccine or as an anticancer agent. The  
 CC present sequence represents a different survivin-derived HLA-A24 antigen,  
 CC survivin 92, which corresponds to residues 115-124 of full-length human

CC survivin (ADA09936) and which was used in an example from the invention.  
 XX  
 SQ Sequence 10 AA;  
 Query Match 7.0%; Score 10; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0025;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 92 QFEELTLGEF 101  
 DB 1 QFEELTLGEF 10  
 RESULT 34  
 AAEL1070  
 ID AAEL1070 standard; peptide; 17 AA.  
 XX  
 AC AAEL1070;  
 XX  
 XX 18-DEC-2001 (first entry)  
 DT Human survivin peptide mutant.  
 XX  
 DE  
 XX Survivin dependent apoptosis; p34(cdc2)-cyclin B1 kinase complex;  
 KW caspase-9; tumour suppressor; antisense therapy; cytostatic; human;  
 KW mutant; mutain.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Modified-site 7..10  
 FT /note= "Phosphorylation site for p34(cdc2)-cyclin B1  
 FT kinase complex"  
 FT Misc-difference 7  
 FT /note= "Wild type Thr substituted with Ala; Corresponds  
 FT to position 34 of wild type survivin peptide"  
 XX  
 XX WO200164741-A2.  
 XX  
 XX 07-SEP-2001.  
 XX  
 XX 28-FEB-2001; 2001WO-US006341.  
 XX  
 XX 29-FEB-2000; 2000US-00515514.  
 XX  
 XX (UYA ) UNIV YALE.  
 XX  
 XX Altieri DC;  
 XX  
 XX WPI; 2001-589861/66.  
 XX  
 XX Modulating survivin apoptosis pathways by using an agent that modulates  
 PT phosphorylation of survivin, interaction between survivin and p34-cyclin  
 PT B1 kinase complex and interaction between survivin and caspase-9.  
 XX  
 PS Example 2; Page; 97pp; English.  
 XX  
 CC The invention relates to a method for modulating apoptosis, in particular  
 CC survivin dependent apoptosis in a cell. The method involves administering  
 CC to the cell an agent which modulates the interaction between survivin and  
 CC p34(cdc2)-cyclin B1 kinase complex, phosphorylation of survivin by the  
 CC p34(cdc2)-cyclin B1 kinase complex and interaction between survivin and  
 CC caspase-9. Survivin antagonist is useful for inhibiting tumour growth in  
 CC a mammal. Agents that modulate, preferably dephosphorylate survivin at  
 CC Thr34, act as potential tumour suppressors and prevent the formation of  
 CC anti-apoptotic complex with caspase. The method is also useful in  
 CC antisense therapy. The present sequence is human survivin peptide mutant  
 CC (T34A). Note: This sequence is not found in the specification but is  
 CC derived from the human survivin peptide sequence (AAEL1066) shown in page  
 CC 45 of the specification  
 XX  
 XX Sequence 17 AA;

Query Match 7.0%; Score 10; DB 4; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 0.0041;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 PERMAEAGFI 44  
 DB 8 PERMAEAGFI 17  
 |||||

RESULT 35  
 AAE11069  
 ID AAE11069 standard; peptide; 9 AA.  
 XX AC AAE11069;  
 XX AC AAE11069;  
 DT 18-DEC-2001 (first entry)  
 XX Human survivin phosphorylation motif.  
 DE  
 XX Survivin dependent apoptosis; p34(cdc2)-cyclin B1 kinase complex;  
 KW caspase-9; tumour suppressor; antisense therapy; cytostatic.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 4..7  
 FT /note= "phosphorylation site for p34(cdc2)-cyclin B1  
 FT kinase complex"  
 FN WO200164741-A2.  
 XX 07-SEP-2001.  
 PD 28-FEB-2001; 2001WO-US006341.  
 XX 29-FEB-2000; 2000US-00515514.  
 FR (UYUA) UNIV YALE.  
 XX PA Altieri DC;  
 XX PI WPI; 2001-589861/66.  
 DR  
 XX Modulating survivin apoptosis pathways by using an agent that modulates  
 PT phosphorylation of survivin, interaction between survivin and p34-cyclin  
 PT B1 kinase complex and interaction between survivin and caspase-9.  
 PS Example 1; Fig 1A; 97pp; English.  
 XX  
 CC The invention relates to a method for modulating apoptosis, in particular  
 CC survivin dependent apoptosis in a cell. The method involves administering  
 CC to the cell an agent which modulates the interaction between survivin and  
 CC p34(cdc2)-cyclin B1 kinase complex, phosphorylation of survivin by the  
 CC p34(cdc2)-cyclin B1 kinase complex and interaction between survivin and  
 CC caspase-9. Survivin antagonist is useful for inhibiting tumour growth in  
 CC a mammal. Agents that modulate, preferably dephosphorylate survivin at  
 CC Thr34, act as potential tumour suppressors and prevent the formation of  
 CC anti-apoptotic complex with caspase. The method is also useful in  
 CC antisense therapy. The present sequence is human survivin phosphorylation  
 CC motif  
 XX Sequence 9 AA;  
 QY 31 CACTPERMA 39  
 DB 1 CACTPERMA 9  
 |||||

Query Match 6.3%; Score 9; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 36  
 ADA09937  
 ID ADA09937 standard; peptide; 9 AA.  
 XX AC ADA09937;  
 XX AC ADA09937;  
 DT 06-NOV-2003 (first entry)  
 XX Human HLA-A24 antigen survivin 85, SEQ ID NO:2.  
 DE  
 XX Human; survivin; survivin 2B80; survivin 85; HLA-A24 antigen;  
 KW human leukocyte antigen; cytotoxic T-cell; CTL; induction; cancer;  
 KW vaccine; anticancer agent; cytostatic.  
 XX OS Homo sapiens.  
 XX OS JP2002284797-A.  
 XX 03-OCT-2002.  
 PD 23-MAR-2001; 2001JP-00084438.  
 XX 23-MAR-2001; 2001JP-00084438.  
 PR (HOKK-) HOKKAIDO TLO KK.  
 XX WPI; 2003-397157/38.  
 DR  
 XX A novel nonapeptide survivin 2B80 capable of inducing cytotoxic T-cells,  
 FT useful in vaccines for the treatment of cancer.  
 FT Example 3; Page 5; 10pp; Japanese.  
 XX  
 CC The invention relates to a human survivin-derived cancer antigen,  
 CC designated survivin 2B80. Survivin 2B80 is a HLA-A24 antigen, and is  
 CC capable of inducing cytotoxic T-cells (CTLs) which target cancer cells  
 CC having HLA-A24 expression. Survivin 2B80 can be used in the treatment of  
 CC cancer, specifically in a cancer vaccine or as an anticancer agent. The  
 CC present sequence represents a different survivin-derived HLA-A24 antigen,  
 CC survivin 85, which corresponds to residues 108-116 of full-length human  
 CC survivin (ADA09936) and which was used in an example from the invention.  
 XX Sequence 9 AA;  
 QY 85 AFLSVKKQF 93  
 DB 1 AFLSVKKQF 9  
 |||||

Query Match 6.3%; Score 9; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 AFLSVKKQF 93  
 DB 1 AFLSVKKQF 9  
 |||||

RESULT 37  
 AAR27745  
 ID AAR27745 standard; protein; 1822 AA.  
 XX AAR27745;  
 XX AAR27745;  
 XX 24-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 04-MAR-1993 (first entry)  
 XX Extracellular factor related protein.  
 DE  
 XX EF\*; detection; prevention; screening; diagnostic.  
 XX Streptococcus suis; type II (non-pathogenic).  
 OS  
 XX Key Location/Qualifiers  
 FT Peptide 1..46  
 FT /note= "signal peptide"  
 FT Peptide 47..1822



Db 3 LGFFLKL 9

RESULT 39  
AAM19295  
ID AAM19295 strand: protein; 73 AA.

AC AAM19295;

DT 12-OCT-2001 (first entry)

DE Peptide #5729 encoded by probe for measuring cervical gene expression.

XX KW Probe: human; microarray; gene expression; cervical epithelial cell;

XX OS cervical cancer.

XX OS Homo sapiens.

XX PN WO200157278-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000670.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX PT gene expression in human cervical epithelial cells.

XX PS Claim 27; SEQ ID NO 24121; 487pp; English.

XX CC The present invention relates to human single exon nucleic acid probes

XX CC (SEN1; see AAM10068-AAM28459). The present sequence is a peptide encoded

XX CC by one such probe. The SENPs are derived from human Hela cells. The SENPs

XX CC can be used to produce a single exon microarray, which can be used for

XX CC measuring human gene expression in a sample derived from human cervical

XX CC epithelial cells. By measuring gene expression, the probes are therefore

XX CC useful in grading and/or staging of diseases of the cervix, notably

XX CC cervical cancer. Note: The sequence data for this patent did not form

XX CC part of the printed specification, but was obtained in electronic format

XX CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 73 AA;

Query Match 4.9%; Score 7; DB 4; Length 73;

Best Local Similarity 100.0%; Pred. No. 25;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 NRIAKET 117

Db 1 NRIAKET 7

RESULT 40

ABB38577

ID ABB38577 standard; peptide; 73 AA.

XX AC ABB38577;

XX DT 04-FEB-2002 (first entry)

XX DE Peptide #6083 encoded by human foetal liver single exon probe.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000669.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483447/52.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX PT gene expression in human fetal liver.

XX PS Claim 27; SEQ ID NO 31212; 639pp + Sequence listing; English.

XX CC The invention relates to a single exon nucleic acid probe for measuring

XX CC human gene expression in a sample derived from human foetal liver. The

XX CC single exon nucleic acid probes may be used for predicting, measuring and

XX CC displaying gene expression in samples derived from human fetal liver. The

XX CC present sequence is a peptide encoded by a single exon nucleic acid probe

XX CC of the invention. Note: The sequence data for this patent did not form

XX CC part of the printed specification, but was obtained in electronic format

XX CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 73 AA;

Query Match 4.9%; Score 7; DB 4; Length 73;

Best Local Similarity 100.0%; Pred. No. 25;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 NRIAKET 117

Db 1 NRIAKET 7

RESULT 41

AAM32028

ID AAM32028 standard; protein; 73 AA.

XX AC AAM32028;

XX DT 17-OCT-2001 (first entry)

XX DE Peptide #6065 encoded by probe for measuring placental gene expression.

XX KW Probe: microarray; human; placenta; antenatal diagnosis;

XX OS genetic disorder.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000663.

PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 XX gene expression in human placenta.  
 XX  
 PS Claim 27; SEQ ID NO 32297; 654bp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SENPs;  
 CC see A133135-A157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders  
 CC  
 SQ Sequence 73 AA;  
 XX  
 Query Match 4.9%; Score 7; DB 4; Length 73;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 111 NRIAKET 117  
 Db 1 NRIAKET 7  
 XX  
 RESULT 42  
 ABB23701  
 ID ABB23701 standard; protein; 73 AA.  
 XX  
 AC ABB23701;  
 XX  
 DT 23-JAN-2002 (first entry)  
 XX  
 DE Protein #5700 encoded by probe for measuring heart cell gene expression.  
 XX  
 KW Human; gene expression; heart; microarray; vascular system;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200157274-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000666.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488897/53.  
 XX

PT Single exon nucleic acid probes for analyzing gene expression in human  
 XX hearts.  
 XX  
 PS Claim 15; SEQ ID NO 25471; 530bp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 73 AA;  
 XX  
 Query Match 4.9%; Score 7; DB 4; Length 73;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 111 NRIAKET 117  
 Db 1 NRIAKET 7  
 XX  
 RESULT 43  
 AAM71736  
 ID AAM71736 standard; protein; 73 AA.  
 XX  
 AC AAM71736;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 32042.  
 XX  
 KW Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukemia; lymphoma; myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200157276-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 XX gene expression in human bone marrow.  
 XX  
 PS Example 4; SEQ ID NO 32042; 658bp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers



CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention

XX  
 SQ Sequence 73 AA;

Query Match 4.9%; Score 7; DB 4; Length 73;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 NKIAKET 117  
 |||||  
 1 NKIAKET 7

RESULT 44  
 AAM59201  
 ID AAM59201 standard; protein; 73 AA.  
 XX  
 AC AAM59201;  
 XX  
 DT 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe encoded protein SEQ ID NO: 31306.

XX Human; brain expressed exon; gene expression analysis; probe; microarray;  
 XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

XX Homo sapiens.

XX WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000667.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WIPI; 2001-483446/52.

XX PT Single exon nucleic acid probes for analyzing gene expression in human

XX PT brains.

XX PS Example 4; SEQ ID NO 31306; 650bp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid

XX CC probes which are derived from genomic sequences expressed in the human

XX CC brain. They can be used to measure gene expression in brain cell samples,

XX CC which may enable the diagnosis and improved treatment of nervous system

XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX CC epilepsy and cancers. The present sequence is a protein encoded by one of

XX CC the probes of the invention

SQ Sequence 73 AA;

Query Match 4.9%; Score 7; DB 4; Length 73;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 NKIAKET 117  
 |||||  
 1 NKIAKET 7

RESULT 45  
 ABG53421  
 ID ABG53421 standard; peptide; 73 AA.  
 XX  
 AC ABG53421;  
 XX  
 DT 25-FEB-2003 (first entry)

XX Human liver peptide, SEQ ID No 32069.

XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

XX hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

XX WO200157273-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000664.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WIPI; 2001-488898/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX PT gene expression in human adult liver.

XX PS Claim 27; SEQ ID NO 32069; 658bp; English.

XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for

XX CC measuring human gene expression in a sample derived from human adult

XX CC liver, comprising one of 13109 defined nucleotide sequences given in the

XX CC specification (or complements/ fragments). The probe hybridises at high

XX CC stringency to a nucleic acid molecule expressed in the human adult liver.

XX CC (I) may be used for predicting, measuring and displaying gene expression

XX CC in samples derived from human adult liver. The genes identified may be

XX CC involved in genetic liver diseases such as cirrhosis,

XX CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is

XX CC associated with coronary heart disease. ABG47348-ABG5930 represent human

XX CC liver single exon encoded peptides of the invention. Note: The sequence

XX CC information for this patent does not appear in the printed specification

XX CC but was obtained in electronic format directly from WIPO at

XX CC ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 73 AA;

Query Match 4.9%; Score 7; DB 4; Length 73;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 NKIAKET 117  
 |||||  
 1 NKIAKET 7

RESULT 46  
 ABG41550  
 ID ABG41550 standard; peptide; 73 AA.  
 XX  
 AC ABG41550;  
 XX  
 DT 19-AUG-2002 (first entry)



CC 06-JUN-1997; 97US-0048972P; 97US-0048974P;  
PR 06-JUN-1997; 97US-0048974P;  
PR 06-JUN-1997; 97US-0049019P;  
PR 06-JUN-1997; 97US-0049020P;  
PR 06-JUN-1997; 97US-0049373P;  
PR 06-JUN-1997; 97US-0049374P;  
PR 06-JUN-1997; 97US-0049375P;  
PR 05-SEP-1997; 97US-005784P;  
PR 05-SEP-1997; 97US-005784P;  
PR 05-SEP-1997; 97US-0057828P;  
PR 05-SEP-1997; 97US-0057829P;  
PR 05-SEP-1997; 97US-0057834P;  
PR 05-SEP-1997; 97US-0057835P;  
PR 05-SEP-1997; 97US-0057842P;  
PR 05-SEP-1997; 97US-0057843P;  
PR 05-SEP-1997; 97US-0057844P;  
PR 05-SEP-1997; 97US-0057845P;  
PR 05-SEP-1997; 97US-0057861P;  
PR 05-SEP-1997; 97US-0057861P;  
PR 05-SEP-1997; 97US-0057862P;  
PR 05-SEP-1997; 97US-0057862P;  
PR 05-SEP-1997; 97US-0057866P;  
PR 05-SEP-1997; 97US-0057867P;  
PR 05-SEP-1997; 97US-0057868P;  
PR 05-SEP-1997; 97US-0057850P;  
PR 05-SEP-1997; 97US-0057851P;  
PR 05-SEP-1997; 97US-0057854P;  
PR 05-SEP-1997; 97US-0057861P;  
PR 05-SEP-1997; 97US-0057862P;  
PR 05-SEP-1997; 97US-0057863P;  
PR 05-SEP-1997; 97US-0057864P;  
PR 05-SEP-1997; 97US-0057865P;  
PR 05-SEP-1997; 97US-0057869P;  
PR 05-SEP-1997; 97US-0057770P;  
PR 05-SEP-1997; 97US-0057771P;  
PR 05-SEP-1997; 97US-0057774P;  
PR 05-SEP-1997; 97US-0057775P;  
PR 05-SEP-1997; 97US-0057776P;  
PR 05-SEP-1997; 97US-0057777P;  
PR 05-SEP-1997; 97US-005778P;  
PR 18-DEC-1997; 97US-0070923P;  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Young P, Greene JW, Ferrie AM, Ruben SM, Rosen CA, Hu J;  
PI Olsen HS, Ehner R, Brewer LA, Moore PA, Shi Y, Florence C;  
PI Florence K, Latteur DW, Ni J, Fan P, Wei Y, Fischer CL, Soppet DR;  
PI Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress G;  
PI Carter KC;  
XX  
DR WPI, 1999-059865/05.  
DR N-PSDB; AAW84612.  
XX  
PT New isolated human genes and the secreted polypeptides they encode -  
XX useful for diagnosis and treatment of e.g. cancers, neurological  
XX disorders, immune diseases, inflammation or blood disorders.  
PS Claim 11, Page 564; 772pp; English.

	CC	determining the presence of mutations in the new polypeptides.
	CC	Specific uses are described for each of the polynucleotides, based on which tissues they are most highly expressed in, and include developing products for the diagnosis or treatment of cancer, neurodegenerative disorders, developmental abnormalities and foetal deficiencies; blood disorders, leucunaz, leukemias, diseases of the immune system, autoimmune diseases, hepatic and renal disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
	CC	restenosis, prostate diseases, obesity, disorders involving osteoclasts such as osteoporosis, arthritis or malfunctions, diseases of testes, lung or thymus, digestive/endocrine disorders, infections and AIDS. The CC polypeptides are also useful for identifying their binding partners. The present sequence represents human secreted protein (see descriptor line for gene number and clone identification)
SQ	XX	Sequence 101 AA;
Query Match		4.9%; Score 7; DB 2; Length 101;
Best Local Similarity		100.0%, Pred. No. 34;
Matches	7;	Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	13 FLDHRI 19       	
D6	22 FLDHRI 28	
RESULT 48		
ID	ABB50502	
AC	ABB50502 standard; protein; 101 AA.	
XX	ABB50502;	
DT	07-FEB-2002 ((first entry))	
DE	Human secreted protein encoded by gene 202 SEQ ID NO:450.	
XX	Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV; dermatological; immunosuppressive; antiinflammatory; immunostimulant; cytotoxic; cardiac; vascular; anti-angiogenic; ophthalmological; neuroprotective; nootropic; anticoagulant; antialzheimers; vulnery; antiParkinsonian; antimicrobial; gene therapy; vaccine; immune disorder; multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer; human immunodeficiency virus; hyperproliferative disorder; wound healing; Gaucher's disease; cardiovascular disease; Schmitz syndrome; chemotaxis; Chagas's cardiomyopathy; coronary arteriosclerosis; angiocentric disorder; cornel graft neovascularisation; diabetic retinopathy; regeneration; cornel grait neovascularisation; Huntington's chorea; Alzheimer's disease; Parkinson's disease; infectious disease.	
KM	Parkinson's disease; infectious disease.	
XX	Homo sapiens.	
OS		
XN	W0200162891-A2.	
PD	30-AUG-2001.	
XX		
PF	21-FEB-2001; 2001WO-US005614.	
PR	24-FEB-2000; 2000US-0184836F.	
FR	29-MAR-2000; 2000US-0193170P.	
PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Nà J., Esher R., Lafleur DW., Moore PA., Olsen HS., Rosen CA.; Ruben SM., Soppe DR., Young PE., Shi Y., Florence KA., Wei Y., Florence C., Hu Y., Kyaw H., Fischer CL., Ferrle AM., Fan P., Feng P., Endress GA., Dillon PJ., Carter KC., Brewer LA., Yu G., Zeng Z., Greene JM;	
WT	WP1: 2001-625724/72. RN-PSDB; ABA83395.	
LT	Nucleic acids encoding 207 human secreted polypeptides, useful for preventing, diagnosing and/or treating, e.g. cancers, Parkinson's disease	

PT and diabetic retinopathy.  
XX  
XX Claim 11, Page 1196, 1533pp; English.  
PS  
CC ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted  
CC proteins (I) and polynucleotide (II) sequences. (I) and (II) have various  
CC activities based on the tissues and cells the genes are expressed in.  
CC Example of these activities include: immunomodulatory; antisclerotic;  
CC dermatological; immunosuppressive; antiinflammatory; immunostimulant;  
CC anti-HIV; cytostatic; cardiant; anti-angiogenic; ophthalmological;  
CC neuroprotective; nootropic; anticonvulsant; antialzheimer's; vascular;  
CC antiparkinsonian; antimicrobial; and vulnerary. (I) and (II) can be used  
CC in gene therapy and vaccine production. (I) and (II) can be used in the  
CC prevention, diagnosis and treatment of immune disorders (e.g. multiple  
CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus  
CC (HIV) infections), hyperproliferative disorders (e.g. cancers and  
CC Gaucher's disease), cardiovascular diseases (e.g. scintar syndrome,  
CC disorders (e.g. corneal graft neovascularization and diabetic  
CC retinopathy), neurological disorders (e.g. Huntington's chorea,  
CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or  
CC for promoting wound healing, regeneration and/or chemotaxis. ABA83195 to  
CC ABA83193 and ABB50300 represent sequences used in the exemplification of  
CC the present invention  
XX  
SQ Sequence 101 AA;  
  
Query Match 4.9%; Score 7; DB 4; Length 101;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 13 FLKDHR 19  
Db 22 FLKDHR 28  
  
RESULT 49  
ABO44759  
ID ABO44759 standard; protein; 101 AA.  
XX  
XX ABO44759;  
AC  
XX  
XX 02-OCT-2003 (first entry)  
DT  
XX  
XX Novel human secreted protein #202.  
DE  
XX  
XX Human; gene therapy; autoimmune disorder; multiple sclerosis; cancer;  
XX systeme lupus erythematosus; haematopoietic cell disorder; allergy;  
XX agammaglobulinemia; ataxia telangiectasia; blood coagulation disorder;  
XX inflammatory condition; ischaemia-reperfusion injury; infectious disease;  
XX hyperproliferative disorder; purpura; viral infection; regeneration;  
XX bacterial infection; ulcer; Alzheimer's disease.  
OS  
XX Homo sapiens.  
XX  
XX US2003065160-A1.  
PN  
XX  
XX 03-APR-2003.  
PD  
XX  
XX 07-DEC-2001; 2001US-00004860.  
PF  
XX  
XX 06-JUN-1997; 97US-0048875P.  
PR 06-JUN-1997; 97US-0048876P.  
PR 06-JUN-1997; 97US-0048877P.  
PR 06-JUN-1997; 97US-0048878P.  
PR 06-JUN-1997; 97US-0048879P.  
PR 06-JUN-1997; 97US-0048880P.  
PR 06-JUN-1997; 97US-0048881P.  
PR 06-JUN-1997; 97US-0048882P.  
PR 06-JUN-1997; 97US-0048883P.  
PR 06-JUN-1997; 97US-0048884P.  
PR 06-JUN-1997; 97US-0048885P.  
PR 06-JUN-1997; 97US-0048892P.

PR 06-JUN-1997; 97US-0048893P.  
PR 06-JUN-1997; 97US-0048894P.  
PR 06-JUN-1997; 97US-0048895P.  
PR 06-JUN-1997; 97US-0048896P.  
PR 06-JUN-1997; 97US-0048897P.  
PR 06-JUN-1997; 97US-0048898P.  
PR 06-JUN-1997; 97US-0048899P.  
PR 06-JUN-1997; 97US-0048900P.  
PR 06-JUN-1997; 97US-0048901P.  
PR 06-JUN-1997; 97US-0048915P.  
PR 06-JUN-1997; 97US-0048916P.  
PR 06-JUN-1997; 97US-0048917P.  
PR 06-JUN-1997; 97US-0048918P.  
PR 06-JUN-1997; 97US-0048919P.  
PR 06-JUN-1997; 97US-0048920P.  
PR 06-JUN-1997; 97US-0048921P.  
PR 06-JUN-1997; 97US-0048922P.  
PR 06-JUN-1997; 97US-0048923P.  
PR 06-JUN-1997; 97US-0048924P.  
PR 06-JUN-1997; 97US-0048925P.  
PR 06-JUN-1997; 97US-0048926P.  
PR 06-JUN-1997; 97US-0048927P.  
PR 06-JUN-1997; 97US-0048928P.  
PR 06-JUN-1997; 97US-0048929P.  
PR 06-JUN-1997; 97US-0048930P.  
PR 06-JUN-1997; 97US-0048931P.  
PR 06-JUN-1997; 97US-0048932P.  
PR 06-JUN-1997; 97US-0048933P.  
PR 06-JUN-1997; 97US-0048934P.  
PR 06-JUN-1997; 97US-0048935P.  
PR 06-JUN-1997; 97US-0048936P.  
PR 06-JUN-1997; 97US-0048937P.  
PR 06-JUN-1997; 97US-0048938P.  
PR 06-JUN-1997; 97US-0048939P.  
PR 06-JUN-1997; 97US-0048940P.  
PR 06-JUN-1997; 97US-0048941P.  
PR 06-JUN-1997; 97US-0048942P.  
PR 06-JUN-1997; 97US-0048943P.  
PR 06-JUN-1997; 97US-0048944P.  
PR 06-JUN-1997; 97US-0048945P.  
PR 06-JUN-1997; 97US-0048946P.  
PR 06-JUN-1997; 97US-0048947P.  
PR 06-JUN-1997; 97US-0048948P.  
PR 06-JUN-1997; 97US-0048949P.  
PR 06-JUN-1997; 97US-0048950P.  
PR 06-JUN-1997; 97US-0048951P.  
PR 06-JUN-1997; 97US-0048952P.  
PR 06-JUN-1997; 97US-0048953P.  
PR 06-JUN-1997; 97US-0048954P.  
PR 06-JUN-1997; 97US-0048955P.  
PR 06-JUN-1997; 97US-0048956P.  
PR 06-JUN-1997; 97US-0048957P.  
PR 06-JUN-1997; 97US-0048958P.  
PR 06-JUN-1997; 97US-0048959P.  
PR 06-JUN-1997; 97US-0048960P.  
PR 06-JUN-1997; 97US-0048961P.  
PR 06-JUN-1997; 97US-0048962P.  
PR 06-JUN-1997; 97US-0048963P.  
PR 06-JUN-1997; 97US-0048964P.  
PR 06-JUN-1997; 97US-0048965P.  
PR 06-JUN-1997; 97US-0048966P.  
PR 06-JUN-1997; 97US-0048967P.  
PR 06-JUN-1997; 97US-0048968P.  
PR 06-JUN-1997; 97US-0048969P.  
PR 06-JUN-1997; 97US-0048970P.  
PR 06-JUN-1997; 97US-0048971P.  
PR 06-JUN-1997; 97US-0048972P.  
PR 06-JUN-1997; 97US-0048973P.  
PR 06-JUN-1997; 97US-0048974P.  
PR 06-JUN-1997; 97US-0048975P.  
PR 06-JUN-1997; 97US-0048976P.  
PR 06-JUN-1997; 97US-0048977P.  
PR 06-JUN-1997; 97US-0048978P.  
PR 06-JUN-1997; 97US-0048979P.  
PR 06-JUN-1997; 97US-0048980P.  
PR 06-JUN-1997; 97US-0048981P.  
PR 06-JUN-1997; 97US-0048982P.  
PR 06-JUN-1997; 97US-0048983P.  
PR 06-JUN-1997; 97US-0048984P.  
PR 06-JUN-1997; 97US-0048985P.  
PR 06-JUN-1997; 97US-0048986P.  
PR 06-JUN-1997; 97US-0048987P.  
PR 06-JUN-1997; 97US-0048988P.  
PR 06-JUN-1997; 97US-0048989P.  
PR 06-JUN-1997; 97US-0048990P.  
PR 06-JUN-1997; 97US-0048991P.  
PR 06-JUN-1997; 97US-0048992P.  
PR 06-JUN-1997; 97US-0048993P.  
PR 06-JUN-1997; 97US-0048994P.  
PR 06-JUN-1997; 97US-0048995P.  
PR 06-JUN-1997; 97US-0048996P.  
PR 06-JUN-1997; 97US-0048997P.  
PR 06-JUN-1997; 97US-0048998P.  
PR 06-JUN-1997; 97US-0048999P.  
PR 06-JUN-1997; 97US-0049000P.  
PR 06-JUN-1997; 97US-0049001P.  
PR 06-JUN-1997; 97US-0049002P.  
PR 06-JUN-1997; 97US-0049003P.  
PR 06-JUN-1997; 97US-0049004P.  
PR 06-JUN-1997; 97US-0049005P.  
PR 06-JUN-1997; 97US-0049006P.  
PR 06-JUN-1997; 97US-0049007P.  
PR 06-JUN-1997; 97US-0049008P.  
PR 06-JUN-1997; 97US-0049009P.  
PR 06-JUN-1997; 97US-0049010P.  
PR 06-JUN-1997; 97US-0049011P.  
PR 06-JUN-1997; 97US-0049012P.  
PR 06-JUN-1997; 97US-0049013P.  
PR 06-JUN-1997; 97US-0049014P.  
PR 06-JUN-1997; 97US-0049015P.  
PR 06-JUN-1997; 97US-0049016P.  
PR 06-JUN-1997; 97US-0049017P.  
PR 06-JUN-1997; 97US-0049018P.  
PR 06-JUN-1997; 97US-0049019P.  
PR 06-JUN-1997; 97US-0049020P.  
PR 06-JUN-1997; 97US-0049021P.  
PR 06-JUN-1997; 97US-0049022P.  
PR 06-JUN-1997; 97US-0049023P.  
PR 06-JUN-1997; 97US-0049024P.  
PR 06-JUN-1997; 97US-0049025P.  
PR 06-JUN-1997; 97US-0049026P.  
PR 06-JUN-1997; 97US-0049027P.  
PR 06-JUN-1997; 97US-0049028P.  
PR 06-JUN-1997; 97US-0049029P.  
PR 06-JUN-1997; 97US-0049030P.  
PR 06-JUN-1997; 97US-0049031P.  
PR 06-JUN-1997; 97US-0049032P.  
PR 06-JUN-1997; 97US-0049033P.  
PR 06-JUN-1997; 97US-0049034P.  
PR 06-JUN-1997; 97US-0049035P.  
PR 06-JUN-1997; 97US-0049036P.  
PR 06-JUN-1997; 97US-0049037P.  
PR 06-JUN-1997; 97US-0049038P.  
PR 06-JUN-1997; 97US-0049039P.  
PR 06-JUN-1997; 97US-0049040P.  
PR 06-JUN-1997; 97US-0049041P.  
PR 06-JUN-1997; 97US-0049042P.  
PR 06-JUN-1997; 97US-0049043P.  
PR 06-JUN-1997; 97US-0049044P.  
PR 06-JUN-1997; 97US-0049045P.  
PR 06-JUN-1997; 97US-0049046P.  
PR 06-JUN-1997; 97US-0049047P.  
PR 06-JUN-1997; 97US-0049048P.  
PR 06-JUN-1997; 97US-0049049P.  
PR 06-JUN-1997; 97US-0049050P.  
PR 06-JUN-1997; 97US-0049051P.  
PR 06-JUN-1997; 97US-0049052P.  
PR 06-JUN-1997; 97US-0049053P.  
PR 06-JUN-1997; 97US-0049054P.  
PR 06-JUN-1997; 97US-0049055P.  
PR 06-JUN-1997; 97US-0049056P.  
PR 06-JUN-1997; 97US-0049057P.  
PR 06-JUN-1997; 97US-0049058P.  
PR 06-JUN-1997; 97US-0049059P.  
PR 06-JUN-1997; 97US-0049060P.  
PR 06-JUN-1997; 97US-0049061P.  
PR 06-JUN-1997; 97US-0049062P.  
PR 06-JUN-1997; 97US-0049063P.  
PR 06-JUN-1997; 97US-0049064P.  
PR 06-JUN-1997; 97US-0049065P.  
PR 06-JUN-1997; 97US-0049066P.  
PR 06-JUN-1997; 97US-0049067P.  
PR 06-JUN-1997; 97US-0049068P.  
PR 06-JUN-1997; 97US-0049069P.  
PR 06-JUN-1997; 97US-0049070P.  
PR 06-JUN-1997; 97US-0049071P.  
PR 06-JUN-1997; 97US-0049072P.  
PR 06-JUN-1997; 97US-0049073P.  
PR 06-JUN-1997; 97US-0049074P.  
PR 06-JUN-1997; 97US-0049075P.  
PR 06-JUN-1997; 97US-0049076P.  
PR 06-JUN-1997; 97US-0049077P.  
PR 06-JUN-1997; 97US-0049078P.  
PR 06-JUN-1997; 97US-0049079P.  
PR 06-JUN-1997; 97US-0049080P.  
PR 06-JUN-1997; 97US-0049081P.  
PR 06-JUN-1997; 97US-0049082P.  
PR 06-JUN-1997; 97US-0049083P.  
PR 06-JUN-1997; 97US-0049084P.  
PR 06-JUN-1997; 97US-0049085P.  
PR 06-JUN-1997; 97US-0049086P.  
PR 06-JUN-1997; 97US-0049087P.  
PR 06-JUN-1997; 97US-0049088P.  
PR 06-JUN-1997; 97US-0049089P.  
PR 06-JUN-1997; 97US-0049090P.  
PR 06-JUN-1997; 97US-0049091P.  
PR 06-JUN-1997; 97US-0049092P.  
PR 06-JUN-1997; 97US-0049093P.  
PR 06-JUN-1997; 97US-0049094P.  
PR 06-JUN-1997; 97US-0049095P.  
PR 06-JUN-1997; 97US-0049096P.  
PR 06-JUN-1997; 97US-0049097P.  
PR 06-JUN-1997; 97US-0049098P.  
PR 06-JUN-1997; 97US-0049099P.  
PR 06-JUN-1997; 97US-0049100P.  
PR 06-JUN-1997; 97US-0049101P.  
PR 06-JUN-1997; 97US-0049102P.  
PR 06-JUN-1997; 97US-0049103P.  
PR 06-JUN-1997; 97US-0049104P.  
PR 06-JUN-1997; 97US-0049105P.  
PR 06-JUN-1997; 97US-0049106P.  
PR 06-JUN-1997; 97US-0049107P.  
PR 06-JUN-1997; 97US-0049108P.  
PR 06-JUN-1997; 97US-0049109P.  
PR 06-JUN-1997; 97US-0049110P.  
PR 06-JUN-1997; 97US-0049111P.  
PR 06-JUN-1997; 97US-0049112P.  
PR 06-JUN-1997; 97US-0049113P.  
PR 06-JUN-1997; 97US-0049114P.  
PR 06-JUN-1997; 97US-0049115P.  
PR 06-JUN-1997; 97US-0049116P.  
PR 06-JUN-1997; 97US-0049117P.  
PR 06-JUN-1997; 97US-0049118P.  
PR 06-JUN-1997; 97US-0049119P.  
PR 06-JUN-1997; 97US-0049120P.  
PR 06-JUN-1997; 97US-0049121P.  
PR 06-JUN-1997; 97US-0049122P.  
PR 06-JUN-1997; 97US-0049123P.  
PR 06-JUN-1997; 97US-0049124P.  
PR 06-JUN-1997; 97US-0049125P.  
PR 06-JUN-1997; 97US-0049126P.  
PR 06-JUN-1997; 97US-0049127P.  
PR 06-JUN-1997; 97US-0049128P.  
PR 06-JUN-1997; 97US-0049129P.  
PR 06-JUN-1997; 97US-0049130P.  
PR 06-JUN-1997; 97US-0049131P.  
PR 06-JUN-1997; 97US-0049132P.  
PR 06-JUN-1997; 97US-0049133P.  
PR 06-JUN-1997; 97US-0049134P.  
PR 06-JUN-1997; 97US-0049135P.  
PR 06-JUN-1997; 97US-0049136P.  
PR 06-JUN-1997; 97US-0049137P.  
PR 06-JUN-1997; 97US-0049138P.  
PR 06-JUN-1997; 97US-0049139P.  
PR 06-JUN-1997; 97US-0049140P.  
PR 06-JUN-1997; 97US-0049141P.  
PR 06-JUN-1997; 97US-0049142P.  
PR 06-JUN-1997; 97US-0049143P.  
PR 06-JUN-1997; 97US-0049144P.  
PR 06-JUN-1997; 97US-0049145P.  
PR 06-JUN-1997; 97US-0049146P.  
PR 06-JUN-1997; 97US-0049147P.  
PR 06-JUN-1997; 97US-0049148P.  
PR 06-JUN-1997; 97US-0049149P.  
PR 06-JUN-1997; 97US-0049150P.  
PR 06-JUN-1997; 97US-0049151P.  
PR 06-JUN-1997; 97US-0049152P.  
PR 06-JUN-1997; 97US-0049153P.  
PR 06-JUN-1997; 97US-0049154P.  
PR 06-JUN-1997; 97US-0049155P.  
PR 06-JUN-1997; 97US-0049156P.  
PR 06-JUN-1997; 97US-0049157P.  
PR 06-JUN-1997; 97US-0049158P.  
PR 06-JUN-1997; 97US-0049159P.  
PR 06-JUN-1997; 97US-0049160P.  
PR 06-JUN-1997; 97US-0049161P.  
PR 06-JUN-1997; 97US-0049162P.  
PR 06-JUN-1997; 97US-0049163P.  
PR 06-JUN-1997; 97US-0049164P.  
PR 06-JUN-1997; 97US-0049165P.  
PR 06-JUN-1997; 97US-0049166P.  
PR 06-JUN-1997; 97US-0049167P.  
PR 06-JUN-1997; 97US-0049168P.  
PR 06-JUN-1997; 97US-0049169P.  
PR 06-JUN-1997; 97US-0049170P.  
PR 06-JUN-1997; 97US-0049171P.  
PR 06-JUN-1997; 97US-0049172P.  
PR 06-JUN-1997; 97US-0049173P.  
PR 06-JUN-1997; 97US-0049174P.  
PR 06-JUN-1997; 97US-0049175P.  
PR 06-JUN-1997; 97US-0049176P.  
PR 06-JUN-1997; 97US-0049177P.  
PR 06-JUN-1997; 97US-0049178P.  
PR 06-JUN-1997; 97US-0049179P.  
PR 06-JUN-1997; 97US-0049180P.  
PR 06-JUN-1997; 97US-0049181P.  
PR 06-JUN-1997; 97US-0049182P.  
PR 06-JUN-1997; 97US-0049183P.  
PR 06-JUN-1997; 97US-0049184P.  
PR 06-JUN-1997; 97US-0049185P.  
PR 06-JUN-1997; 97US-0049186P.  
PR 06-JUN-1997; 97US-0049187P.  
PR 06-JUN-1997; 97US-0049188P.  
PR 06-JUN-1997; 97US-0049189P.  
PR 06-JUN-1997; 97US-0049190P.  
PR 06-JUN-1997; 97US-0049191P.  
PR 06-JUN-1997; 97US-0049192P.  
PR 06-JUN-1997; 97US-0049193P.  
PR 06-JUN-1997; 97US-0049194P.  
PR 06-JUN-1997; 97US-0049195P.  
PR 06-JUN-1997; 97US-0049196P.  
PR 06-JUN-1997; 97US-0049197P.  
PR 06-JUN-1997; 97US-0049198P.  
PR 06-JUN-1997; 97US-0049199P.  
PR 06-JUN-1997; 97US-0049200P.  
PR 06-JUN-1997; 97US-0049201P.  
PR 06-JUN-1997; 97US-0049202P.  
PR 06-JUN-1997; 97US-0049203P.  
PR 06-JUN-1997; 97US-0049204P.  
PR 06-JUN-1997; 97US-0049205P.  
PR 06-JUN-1997; 97US-0049206P.  
PR 06-JUN-1997; 97US-0049207P.  
PR 06-JUN-1997; 97US-0049208P.  
PR 06-JUN-1997; 97US-0049209P.  
PR 06-JUN-1997; 97US-0049210P.  
PR 06-JUN-1997; 97US-0049211P.  
PR 06-JUN-1997; 97US-0049212P.  
PR 06-JUN-1997; 97US-0049213P.  
PR 06-JUN-1997; 97US-0049214P.  
PR 06-JUN-1997; 97US-0049215P.  
PR 06-JUN-1997; 97US-0049216P.  
PR 06-JUN-1997; 97US-0049217P.  
PR 06-JUN-1997; 97US-0049218P.  
PR 06-JUN-1997; 97US-0049219P.  
PR 06-JUN-1997; 97US-0049220P.  
PR 06-JUN-1997; 97US-0049221P.  
PR 06-JUN-1997; 97US-0049222P.  
PR 06-JUN-1997; 97US-0049223P.  
PR 06-JUN-1997; 97US-0049224P.  
PR 06-JUN-1997; 97US-0049225P.  
PR 06-JUN-1997; 97US-0049226P.  
PR 06-JUN-1997; 97US-0049227P.  
PR 06-JUN-1997; 97US-0049228P.  
PR 06-JUN-1997; 97US-0049229P.  
PR 06-JUN-1997; 97US-0049230P.  
PR 06-JUN-1997; 97US-0049231P.  
PR 06-JUN-1997; 97US-0049232P.  
PR 06-JUN-1997; 97US-0049233P.  
PR 06-JUN-1997; 97US-0049234P.  
PR 06-JUN-1997; 97US-0049235P.  
PR 06-JUN-1997; 97US-0049236P.  
PR 06-JUN-1997; 97US-0049237P.  
PR 06-JUN-1997; 97US-0049238P.  
PR 06-JUN-1997; 97US-0049239P.  
PR 06-JUN-1997; 97US-0049240P.  
PR 06-JUN-1997; 97US-0049241P.  
PR 06-JUN-1997; 97US-0049242P.  
PR 06-JUN-1997; 97US-0049243P.  
PR 06-JUN-1997; 97US-0049244P.  
PR 06-JUN-1997; 97US-0049245P.  
PR 06-JUN-1997; 97US-0049246P.  
PR 06-JUN-1997; 97US-0049247P.  
PR 06-JUN-1997; 97US-0049248P.  
PR 06-JUN-1997; 97US-0049249P.  
PR 06-JUN-1997; 97US-0049250P.  
PR 06-JUN-1997; 97US-0049251P.  
PR 06-JUN-1997; 97US-0049252P.  
PR 06-JUN-1997; 97US-0049253P.  
PR 06-JUN-1997; 97US-0049254P.  
PR 06-JUN-1997; 97US-0049255P.  
PR 06-JUN-1997; 97US-0049256P.  
PR 06-JUN-1997; 97US-0049257P.  
PR 06-JUN-1997; 97US-0049258P.  
PR 06-JUN-1997; 97US-0049259P.  
PR 06-JUN-1997; 97US-0049260P.  
PR 06-JUN-1997; 97US-0049261P.  
PR 06-JUN-1997; 97US-0049262P.  
PR 06-JUN-1997; 97US-0049263P.  
PR 06-JUN-1997; 97US-0049264P.  
PR 06-JUN-1997; 97US-0049265P.  
PR 06-JUN-1997; 97US-0049266P.  
PR 06-JUN-1997; 97US-0049267P.  
PR 06-JUN-1997; 97US-0049268P.  
PR 06-JUN-1997; 97US-0049269P.  
PR 06-JUN-1997; 97US-0049270P.  
PR 06-JUN-1997; 97US-0049271P.  
PR 06-JUN-1997; 97US-0049272P.  
PR 06-JUN-1997; 97US-0049273P.  
PR 06-JUN-1997; 97US-0049274P.  
PR 06-JUN-1997; 97US-0049275P.  
PR 06-JUN-1997; 97US-0049276P.  
PR 06-JUN-1997; 97US-0049277P.  
PR 06-JUN-1997; 97US-0049278P.  
PR 06-JUN-1997; 97US-0049279P.  
PR 06-JUN-1997; 97US-0049280P.  
PR 06-JUN-1997; 97US-0049281P.  
PR 06-JUN-1997; 97US-0049282P.  
PR 06-JUN-1997; 97US-0049283P.  
PR 06-JUN-1997; 97US-0049284P.  
PR 06-JUN-1997; 97US-0049285P.  
PR 06-JUN-1997; 97US-0049286P.  
PR 06-JUN-1997; 97US-0049287P.  
PR 06-JUN-1997; 97US-0049288P.  
PR 06-JUN-1997; 97US-0049289P.  
PR 06-JUN-1997; 97US-0049290P.  
PR 06-JUN-1997; 97US-0049291P.  
PR 06-JUN-1997; 97US-0049292P.  
PR 06-JUN-1997; 97US-0049293P.  
PR 06-JUN-1997; 97US-0049294P.  
PR 06-JUN-1997; 97US-0049295P.  
PR 06-JUN-1997; 97US-0049296P.  
PR 06-JUN-1997; 97US-0049297P.  
PR 06-JUN-1997; 97US-0049298P.  
PR 06-JUN-

PI Carter KC;  
XX WPI, 2003-540804/51.  
DR N-PSDB; ACH04896.  
XX  
PT New isolated protein, useful for preparing a composition for diagnosing  
PT or treating cancer, inflammatory, immune or infectious diseases.  
XX  
PS Disclosure; SEQ ID NO 450; 1722p; English.  
XX  
XX The invention relates to an isolated HEMA80 protein. The protein is  
CC useful for preparing a composition for diagnosing or treating autoimmune  
CC disorders e.g. multiple sclerosis and systemic lupus erythematosus;  
CC haematopoietic cell disorders e.g. agammaglobulinaemia and ataxia  
CC telangiectasia; blood coagulation disorders e.g. afibrinogenemia and  
CC thrombocytopenia; allergy; graft-versus-host disease; inflammatory  
CC conditions e.g. ischaemia-reperfusion injury and arthritis;  
CC hyperproliferative disorders e.g. cancer and psoriasis; infectious disease  
CC e.g. viral infection and bacterial infection. The polynucleotide or  
CC protein can be used to regenerate damaged tissue e.g. ulcers and  
CC Alzheimer's disease. The present sequence represents the amino acid  
CC sequence of a novel human secreted protein. Note: The sequence data for  
CC this patent did not form part of the printed specification but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030065160  
XX  
SQ Sequence 101 AA;  
Query Match 4.9%; Score 7; DB 6; Length 101;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 13 FLKDHR1 19  
Db 22 FLKDHR1 28  
RESULT 50  
ABO26239  
ID ABO26239 standard; protein; 101 AA.  
XX  
AC ABO26239;  
XX  
DT 10-SEP-2003 (first entry)  
XX  
DE Human protein from novel secreted protein gene 202.  
XX  
KW Human; secreted protein; precerebellin-like protein;  
KW neurodegenerative disorder; behavioural disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; schizophrenia; mania;  
KW dementia; paranoia; psychosis; autism; immune disorder; infection;  
KW inflammation; allergy; liver disorder; hepatoblastoma; jaundice;  
KW hepatitis; immunological disorder; AIDS; leukaemia; rheumatoid arthritis;  
KW sepsis; acne; psoriasis; cancer.  
XX  
XX Homo sapiens.  
XX  
FN US6525174-B1.  
XX  
PD 25-FEB-2003.  
XX  
XX 04-DEC-1998; 98US-00205258.  
XX  
XX 06-JUN-1997; 97US-0048875P.  
PR 06-JUN-1997; 97US-0048876P.  
PR 06-JUN-1997; 97US-0048877P.  
PR 06-JUN-1997; 97US-0048878P.  
PR 06-JUN-1997; 97US-0048880P.  
PR 06-JUN-1997; 97US-0048881P.  
PR 06-JUN-1997; 97US-0048882P.  
PR 06-JUN-1997; 97US-0048883P.  
PR 06-JUN-1997; 97US-0048884P.  
PR 06-JUN-1997; 97US-0048885P.

PR 06-JUN-1997; 97US-0048882P.  
PR 06-JUN-1997; 97US-0048883P.  
PR 06-JUN-1997; 97US-0048884P.  
PR 06-JUN-1997; 97US-0048885P.  
PR 06-JUN-1997; 97US-0048886P.  
PR 06-JUN-1997; 97US-0048887P.  
PR 06-JUN-1997; 97US-0048888P.  
PR 06-JUN-1997; 97US-0048889P.  
PR 06-JUN-1997; 97US-0048900P.  
PR 06-JUN-1997; 97US-0048901P.  
PR 06-JUN-1997; 97US-0048915P.  
PR 06-JUN-1997; 97US-0048916P.  
PR 06-JUN-1997; 97US-0048917P.  
PR 06-JUN-1997; 97US-0048929P.  
PR 06-JUN-1997; 97US-0048933P.  
PR 06-JUN-1997; 97US-0048964P.  
PR 06-JUN-1997; 97US-0048970P.  
PR 06-JUN-1997; 97US-0048971P.  
PR 06-JUN-1997; 97US-0048972P.  
PR 06-JUN-1997; 97US-0048974P.  
PR 06-JUN-1997; 97US-0049019P.  
PR 06-JUN-1997; 97US-0049020P.  
PR 06-JUN-1997; 97US-0049373P.  
PR 06-JUN-1997; 97US-0049374P.  
PR 06-JUN-1997; 97US-0049375P.  
PR 05-SEP-1997; 97US-0057564P.  
PR 05-SEP-1997; 97US-0057628P.  
PR 05-SEP-1997; 97US-0057629P.  
PR 05-SEP-1997; 97US-0057630P.  
PR 05-SEP-1997; 97US-0057631P.  
PR 05-SEP-1997; 97US-0057632P.  
PR 05-SEP-1997; 97US-0057633P.  
PR 05-SEP-1997; 97US-0057634P.  
PR 05-SEP-1997; 97US-0057643P.  
PR 05-SEP-1997; 97US-0057644P.  
PR 05-SEP-1997; 97US-0057645P.  
PR 05-SEP-1997; 97US-0057646P.  
PR 05-SEP-1997; 97US-0057647P.  
PR 05-SEP-1997; 97US-0057648P.  
PR 05-SEP-1997; 97US-0057649P.  
PR 05-SEP-1997; 97US-0057650P.  
PR 05-SEP-1997; 97US-0057651P.  
PR 05-SEP-1997; 97US-0057654P.  
PR 05-SEP-1997; 97US-0057661P.  
PR 05-SEP-1997; 97US-0057662P.  
PR 05-SEP-1997; 97US-0057666P.  
PR 05-SEP-1997; 97US-0057667P.  
PR 05-SEP-1997; 97US-0057668P.  
PR 05-SEP-1997; 97US-0057760P.  
PR 05-SEP-1997; 97US-0057761P.  
PR 05-SEP-1997; 97US-0057762P.  
PR 05-SEP-1997; 97US-0057763P.  
PR 05-SEP-1997; 97US-0057764P.  
PR 05-SEP-1997; 97US-0057765P.  
PR 05-SEP-1997; 97US-0057766P.  
PR 05-SEP-1997; 97US-0057767P.  
PR 05-SEP-1997; 97US-0057778P.  
PR 05-SEP-1997; 97US-0070923P.  
PR 04-DEC-1997; 98WO-US011422.  
PR 15-JUL-1998; 98US-0092521P.  
PR 30-JUN-1998; 98US-0094657P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Young P, Greene JV, Ferrie AM, Ruben SM, Rosen CA, Hu J;  
PI Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y, Florence C;  
PI Florence K, Lafleur DW, Ni J, Fan P, Wei Y, Fischer CL, Soppet DR,  
Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress GA;

PI Carter KC;  
 XX WPI: 2003-511926/48.  
 DR N-FSD; AC044706.  
 XX  
 PT New preserebellin-like protein, useful for diagnosing or treating  
 PT neurodegenerative and behavioral disorders, immune disorders, liver  
 PT disorders, and cancer.  
 XX  
 PS Disclosure; SEQ ID NO 450; 156pp; English.  
 XX  
 CC The invention relates to an isolated protein comprising amino acid  
 CC residues 33-205 or 1-205 of a novel human secreted protein appearing as  
 CC ABO26252. The protein is encoded by one of 238 disclosed cDNA sequences  
 CC encoding 238 secreted proteins. ABO26252 is a preserebellin-like protein.  
 CC Also included are a composition comprising the protein and a carrier and  
 CC an isolated protein produced by expressing the protein cited above by a  
 CC cell, and recovering the protein. The proteins are useful for diagnosing  
 CC or treating neurodegenerative and behavioral disorders (e.g. Alzheimer's  
 CC disease, Parkinson's disease, Huntington's disease, schizophrenia, mania,  
 CC dementia, paranoia, psychoses or autism), immune disorders (e.g.  
 CC infection, inflammation, allergy), liver disorders (e.g. hepatoblastoma,  
 CC jaundice, hepatitis), immunological disorders (e.g. AIDS, leukaemia,  
 CC rheumatoid arthritis, sepsis, acne, psoriasis) and cancer. The present  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from USPTO  
 CC at: [seqdata.uspto.gov/sequence.html?docid=6525174B1](http://seqdata.uspto.gov/sequence.html?docid=6525174B1)  
 CC  
 SQ Sequence 101 AA;  
 XX  
 QY  
 DB 13 FLKDHRI 19  
 22 FLKDHRI 28  
 XX  
 RESULT 51  
 AD088028  
 ID AD088028 standard; protein; 130 AA.  
 XX  
 AC AD088028;  
 XX  
 DT 01-JAN-2004 (first entry)  
 XX  
 DS Ribosomal protein similar to FCWPI #244.  
 XX  
 XX Antifungal protein; ribosomal protein; FCWPI; ALYAF;  
 KW plant fungal infection; Alternaria; Ascochyta; Botrytis; Cercospora;  
 KW Colletotrichum; Diploidea; Fusarium; Gaeanomyces; Helminthosporium;  
 KW Macrophomina; Mycosphaerella; Nectria; Peronospora; Phoma;  
 KW Phymatotrichum; Phytophthora; Plasmopara; Podospaera; Puccinia; Puthium;  
 KW Pyrenophora; Pyricularia; Pythium; Rhizoctonia; Sclerotium; Sclerotinia;  
 KW Septoria; Thielaviopsis; Venturia; Verticillium.  
 XX  
 OS Unidentified.  
 XX  
 PN US6573361-B1.  
 XX  
 PD 03-JUN-2003.  
 XX  
 PF 07-DEC-2000; 2000US-00732210.  
 XX  
 PR 07-DEC-1999; 99US-0169340P.  
 PR 07-DEC-1999; 99US-0169513P.  
 XX  
 PA (MONS ) MONSANTO TECHNOLOGY LLC.  
 XX  
 PI Bunkers GU, Liang J, Miltanck CA, Seale JW, Wu YS;

DR WPI: 2003-754556/71.  
 XX  
 PT Novel antifungal protein FCWPI, isolated from *Fusarium culmorum*, useful  
 PT for controlling fungal infections in plants.  
 XX  
 PS Example 21; SEQ ID NO 281; 27pp; English.  
 XX  
 CC The invention relates to an isolated antifungal ribosomal protein from  
 CC *Fusarium culmorum*, FCWPI. Also included is a fusion protein between the  
 CC signal peptide of the antifungal protein ALYAF from *Alyseum* and FCWPI,  
 CC encoded by the nucleic acid appearing as AD087758. The FCWPI proteins are  
 CC useful for controlling fungal infections in plants, such as those caused  
 CC by *Alternaria* (e.g. *Alternaria brassicae*, *Alternaria solani*),  
 CC *Ascochyta* (e.g. *Ascochyta pisi*), *Botrytis* (e.g. *Botrytis cinerea*),  
 CC *Cercospora* (e.g. *Cercospora kikuchii*, *Cercospora zea-maydis*),  
 CC *Colletotrichum* (e.g. *Colletotrichum lindemuthianum*), *Diploidea* (e.g.  
 CC *Diploidea maydis*), *Fusarium* (e.g. *Fusarium nivale*, *Fusarium oxysporum*,  
 CC *Fusarium graminearum*, *Fusarium culmorum*, *Fusarium solani*, *Fusarium*  
 CC *moniliforme*, *Fusarium roseum*), *Gaeumannomyces* (e.g. *Gaeumannomyces*  
 CC *graminis* f.sp. *tritici*), *Helminthosporium* (e.g. *Helminthosporium turcicum*  
 CC), *Helminthosporium carbonum*, *Helminthosporium maydis*, *Macrophomina*  
 CC (e.g. *Macrophomina phaseolina*, *Magnaporthe grisea*), *Mycosphaerella*  
 CC (e.g. *Mycosphaerella figiensis*), *Nectria* (*Nectria haematococca*),  
 CC *Peronospora* (e.g. *Peronospora manshurica*, *Peronospora tabacina*), *Phoma*  
 CC (e.g. *Phoma betae*), *Phymatotrichum* (e.g. *Phymatotrichum omnivorum*),  
 CC *Phytophthora* (e.g. *Phytophthora cinnamomi*, *Phytophthora cactorum*,  
 CC *Phytophthora phaseoli*, *Phytophthora parasitica*, *Phytophthora*  
 CC *citrophthora*, *Phytophthora megasperma* f.sp. *sojae*, *Phytophthora*  
 CC *infestans*), *Plasmopara* (e.g. *Plasmopara viticola*), *Podospaera* (e.g.  
 CC *Podospaera leucotricha*), *Puccinia* (e.g. *Puccinia sorghi*, *Puccinia*  
 CC *stiriformis*, *Puccinia graminis* f.sp. *tritici*, *Puccinia asparagi*,  
 CC *Puccinia recondita*, *Puccinia arachidis*), *Puthium* (e.g. *Puthium*  
 CC *aphanidermatum*), *Pyrenophora* (e.g. *Pyrenophora tritici-repentens*), (e.g.  
 CC *Pyricularia* (e.g. *Pyricularia oryzae*), *Pythium* (e.g. *Pythium ultimum*),  
 CC *Rhizoctonia* (e.g. *Rhizoctonia solani*, *Rhizoctonia cerealis*), *Sclerotium*  
 CC (e.g. *Sclerotium rolfsii*), *Sclerotinia* (e.g. *Sclerotinia sclerotiorum*),  
 CC *Septoria* (e.g. *Septoria lycopersici*, *Septoria glycines*, *Stagonospora*  
 CC *nodorum* / *Phaeosphaeria nodorum*, *Septoria tritici*), *Thielaviopsis* (e.g.  
 CC *Thielaviopsis basicola*), *Uncinula* (e.g. *Uncinula necator*), *Venturia*  
 CC (e.g. *Venturia inaequalis*) or *Verticillium* (e.g. *Verticillium dahliae*,  
 CC *Verticillium albo-atrum*). Mutations in the proteolytic consensus  
 CC sequences contained within FCWPI provides improved stability of its  
 CC antifungal activity. Also disclosed are ribosomal proteins with similar  
 CC PI (>71) and molecular weight (<20kDa) to FCWPI, which may act as  
 CC antifungal proteins. The present sequence represents one of the ribosomal  
 CC proteins similar to FCWPI. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from USPTO at  
 CC [seqdata.uspto.gov/sequence.html?docid=6573361B1](http://seqdata.uspto.gov/sequence.html?docid=6573361B1).  
 CC  
 SQ Sequence 130 AA;  
 XX  
 QY  
 DB 94 EELTIGE 100  
 47 EELTIGE 53  
 XX  
 RESULT 52  
 AD095705  
 ID AD095705 standard; protein; 167 AA.  
 XX  
 AC AD095705;  
 XX  
 DT 01-JAN-2004 (first entry)  
 XX  
 DS E. faecium protein sequence SEQ ID 5332.  
 XX  
 PA Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;  
 XX abdominal-pelvic infection.  
 KW

```

XX OS Enterococcus faecium.
XX PN US6583275-B1.
XX PD 24-JUN-2003.
XX PF 30-JUN-1998; 98US-00107532.
XX PR 02-JUL-1997; 97US-0051571P.
XX PR 14-MAY-1998; 98US-0085598P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Doucette-Stamm LA, Bush D;
XX DR WPI; 2003-799836/75.
XX DR N-PSDB; ADC92051.
XX PT New isolated nucleic acid derived from Enterococcus faecium encoding an
XX PT Enterococcus faecium polypeptide useful for detection, prevention and
XX PT treatment of a pathological condition resulting from a bacterial
XX PT infection.
XX PS Example 1; SEQ ID NO 5332; 243pp; English.
XX CC The invention relates to an isolated nucleic acid derived from
XX CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
XX CC one of 10 fully defined sequences given in the (or comprising 40
XX CC sequential nucleotides chosen from any of the nucleic acids, its
XX CC complement or sequences hybridizing to it). Also included are a
XX CC recombinant vector comprising the nucleic acid operably linked to a
XX CC transcription regulatory element, a cell comprising the vector and a
XX CC single-stranded probe comprising the nucleic acid. The nucleic acids are
XX CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
XX CC The nucleic acids is useful for diagnosing pathological conditions
XX CC resulting from E. faecium bacterial infection (e.g. urinary tract
XX CC infection, bacteraemia, endocarditis, wounds and abdominal pelvic
XX CC infection) and for screening drugs such as agonists and antagonists. The
XX CC nucleic acid is useful for recombinant production of Candida albicans -
XX CC derived peptides or antisense polypeptides. Pharmaceutical compositions
XX CC and vaccines containing the nucleic acid are useful for preventing or
XX CC treating Enterococcus faecium infections. The present sequence represents
XX CC one if the disclosed E. faecium proteins.
XX SQ Sequence 167 AA;
XX
XX Query Match 4.9%; Score 7; DB 7; Length 167;
XX Best Local Similarity 100.0%; Pred. No. 54;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Oy 131 VRRATFQ 137
XX |||||
XX 145 VRRATFQ 151
XX
XX Db
XX
XX RESULT 53
XX ABB48783
XX ID ABB48783 standard; protein; 179 AA.
XX AC ABB48783;
XX DT 05-FEB-2002 (first entry)
XX DE Listeria monocytogenes protein #1487.
XX KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX KM vitamin B12; bacterial infection; disease.
XX OS Listeria monocytogenes.
XX OS WO200177335-A2.
XX PN
XX XX

```

```

PD 18-OCT-2001.
XX XX
XX PF 11-APR-2001; 2001NO-FR001118.
XX XX
XX PR 11-APR-2000; 2000FR-00004629.
XX XX
XX PA (INSP ) INST PASTEUR.
XX
XX PI Buchrieser C, Franceni L, Couve E, Rusnok C, Fehli H, Dehoux P;
XX PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
XX PI Daniels J, Goebel W, Kieft U, Kuhn M, Ng E, Vazquez-Boland JA;
XX PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
XX PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
XX PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
XX PI Madueno B, De Pablos B, Wehland J, Kaerst U, Ertian K, Hauf U;
XX PI Rose M, Voss H;
XX DR WPI; 2002-010914/01.
XX DR
XX PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX PT and prevention of Listeria and related bacterial infections, and related
XX PT polypeptides.
XX PS Claim 6; SEQ ID NO 1488; 192pp; French.
XX
XX CC The present invention relates to the genome sequence of Listeria
XX CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
XX CC it are useful for selecting probes and primers for detecting genes in L.
XX CC monocytogenes and related organisms, and for studying genetic
XX CC polymorphisms and other genomes. The present sequence is a protein
XX CC encoded by the genome sequence of the present invention. Proteins
XX CC expressed from the genome sequence are useful for raising specific
XX CC antibodies, identification of L. monocytogenes and related organisms, and
XX CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
XX CC B12. The genome sequence and proteins encoded by it are also useful for
XX CC selecting compounds that regulate gene expression and cell replication
XX CC and modulate L. monocytogenes-related diseases. In addition, the genome
XX CC sequence and proteins encoded by it are useful in pharmaceutical and
XX CC vaccines compositions for the treatment or prevention of infections by L.
XX CC monocytogenes and related organisms. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 179 AA;
XX
XX Query Match 4.9%; Score 7; DB 5; Length 179;
XX Best Local Similarity 100.0%; Pred. No. 57;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Oy 127 TAKXVRR 133
XX |||||
XX 36 TAKXVRR 42
XX
XX Db
XX
XX RESULT 54
XX ABU33017
XX ID ABU33017 standard; protein; 179 AA.
XX AC ABU33017;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by prokaryotic essential gene #18544.
XX KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX KM Listeria monocytogenes.
XX OS WO200277183-A2.
XX OS 03-OCT-2002.
XX PN
XX XX

```

21-MAR-2002; 2002WO-US009107.  
21-MAR-2001; 2001US-00815242.  
06-SEP-2001; 2001US-00948993.  
25-OCT-2001; 2001US-0342923P.  
08-FEB-2002; 2002US-00072851.  
06-MAR-2002; 2002US-0362699P.  
(ELIT-) ELITRA PHARM INC.  
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
WPI; 2003-029926/02.  
N-PSDB; ACA36887.  
New antisense nucleic acids, useful for identifying proteins or screening  
for homologous nucleic acids required for cellular proliferation to  
isolate candidate molecules for rational drug discovery programs.  
Claim 25; SEQ ID NO 60941; 1766bp; English.  
The invention relates to an isolated nucleic acid comprising any one of  
the 6213 antisense sequences given in the specification where expression  
of the nucleic acid inhibits proliferation of a cell. Also included are:  
(1) a vector comprising a promoter operably linked to the nucleic acid  
encoding a polypeptide whose expression is inhibited by the antisense  
nucleic acid; (2) a host cell containing the vector; (3) an isolated  
polypeptide or its fragment whose expression is inhibited by the  
antisense nucleic acid; (4) an antibody capable of specifically binding  
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
proliferation or the activity of a gene in an operon required for  
proliferation; (7) identifying a compound that influences the activity of  
the gene product or that has an activity against a biological pathway  
required for proliferation, or that inhibits cellular proliferation; (8)  
identifying a gene required for cellular proliferation or the biological  
pathway in which a proliferation-regulated gene or its gene product lies  
or a gene on which the test compound that inhibits proliferation of an  
organism acts; (9) manufacturing an antibiotic; (10) profiling a  
compound's activity; (11) a culture comprising strains in which the gene  
product is overexpressed or underexpressed; (12) determining the extent  
to which each of the strains is present in a culture or collection of  
strains; or (13) identifying the target of a compound that inhibits the  
proliferation of an organism. The antisense nucleic acids are useful for  
identifying proteins or screening for homologous nucleic acids required  
for cellular proliferation or screening for homologous nucleic acids  
required for cellular proliferation to isolate candidate molecules for rational  
drug discovery programs, or for screening homologous nucleic acids  
required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
*K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
the target prokaryotic essential genes. Note: The sequence data for this  
patent did not form part of the printed specification, but was obtained  
in electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences  
Sequence 179 AA;  
Query Match 4.9%; Score 7; DB 6; Length 179;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
127 TAKKVR 133  
| | | | |  
36 TAKKVR 42  
Db  
RESULT 55  
ABU29201  
ID ABU29201 standard; protein; 182 AA.  
XX AC ABU29201;  
XX DT 19-JUN-2003 (first entry)  
XX

Protein encoded by Prokaryotic essential gene #14728.  
Antisense; prokaryotic essential gene; cell proliferation; drug design.  
*Enterococcus faecalis*.  
WO200277183-A2.  
03-OCT-2002.  
21-MAR-2002; 2002WO-US009107.  
21-MAR-2001; 2001US-00815242.  
06-SEP-2001; 2001US-00948993.  
25-OCT-2001; 2001US-0342923P.  
08-FEB-2002; 2002US-00072851.  
06-MAR-2002; 2002US-0362699P.  
(ELIT-) ELITRA PHARM INC.  
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
WPI; 2003-029926/02.  
N-PSDB; ACA33071.  
New antisense nucleic acids, useful for identifying proteins or screening  
for homologous nucleic acids required for cellular proliferation to  
isolate candidate molecules for rational drug discovery programs.  
Claim 25; SEQ ID NO 57125; 1766bp; English.  
The invention relates to an isolated nucleic acid comprising any one of  
the 6213 antisense sequences given in the specification where expression  
of the nucleic acid inhibits proliferation of a cell. Also included are:  
(1) a vector comprising a promoter operably linked to the nucleic acid  
encoding a polypeptide whose expression is inhibited by the antisense  
nucleic acid; (2) a host cell containing the vector; (3) an isolated  
polypeptide or its fragment whose expression is inhibited by the  
antisense nucleic acid; (4) an antibody capable of specifically binding  
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
proliferation or the activity of a gene in an operon required for  
proliferation; (7) identifying a compound that influences the activity of  
the gene product or that has an activity against a biological pathway  
required for proliferation, or that inhibits cellular proliferation; (8)  
identifying a gene required for cellular proliferation or the biological  
pathway in which a proliferation-regulated gene or its gene product lies  
or a gene on which the test compound that inhibits proliferation of an  
organism acts; (9) manufacturing an antibiotic; (10) profiling a  
compound's activity; (11) a culture comprising strains in which the gene  
product is overexpressed or underexpressed; (12) determining the extent  
to which each of the strains is present in a culture or collection of  
strains; or (13) identifying the target of a compound that inhibits the  
proliferation of an organism. The antisense nucleic acids are useful for  
identifying proteins or screening for homologous nucleic acids required  
for cellular proliferation to isolate candidate molecules for rational  
drug discovery programs, or for screening homologous nucleic acids  
required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
*K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
the target prokaryotic essential genes. Note: The sequence data for this  
patent did not form part of the printed specification, but was obtained  
in electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences  
Sequence 182 AA;  
Query Match 4.9%; Score 7; DB 6; Length 182;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
127 TAKKVR 133  
| | | | |  
39 TAKKVR 45  
Db



RESULT 56  
 ABU29837  
 ID ABU29837 standard; protein; 182 AA.  
 XX  
 XX  
 AC ABU29837;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Protein encoded by Prokaryotic essential gene #15364.  
 XX  
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 OS Enterococcus faecium.  
 XX  
 PN MO200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002MO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trivick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR N-PSDB; ACA33707.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 PS Claim 25; SEQ ID NO 57761; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway; (8)  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs; or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
 SQ Sequence 182 AA;  
 XX  
 Query Match 4.9%; Score 7; DB 6; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 127 TAKYRR 133  
 Db 39 TAKYRR 45  
 RESULT 57  
 ADC95400  
 ID ADC95400 standard; protein; 186 AA.  
 XX  
 AC ADC95400;  
 XX  
 DT 01-JAN-2004 (first entry)  
 XX  
 DE E. faecium protein sequence SEQ ID 5027.  
 XX  
 KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;  
 KW abdominal/pelvic infection.  
 XX  
 OS Enterococcus faecium.  
 XX  
 PN US6583275-B1.  
 XX  
 PD 24-JUN-2003.  
 XX  
 PF 30-JUN-1998; 98US-00107532.  
 XX  
 PR 02-JUL-1997; 97US-0051571P.  
 PR 14-MAY-1998; 98US-0085598P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Doucette-Stamm LA, Bush D;  
 PI  
 XX  
 DR WPI: 2003-799836/75.  
 DR N-PSDB; ADC91746.  
 XX  
 PT New isolated nucleic acid derived from *Enterococcus faecium* encoding an  
 PT *Enterococcus faecium* polypeptide useful for detection, prevention and  
 PT treatment of a pathological condition resulting from a bacterial  
 PT infection.  
 XX  
 PS Example 1; SEQ ID NO 5027; 243pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid derived from  
 CC *Enterococcus faecium* encoding an *Enterococcus faecium* polypeptide having  
 CC one of 10 fully defined sequences given in the (or comprising 40  
 CC sequential nucleotides chosen from any of the nucleic acids, its  
 CC complement or sequences hybridising to it). Also included are a  
 CC recombinant vector comprising the nucleic acid operably linked to  
 CC transcription regulatory element, a cell comprising the vector and a  
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are  
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.  
 CC The nucleic acids is useful for diagnosing pathological conditions  
 CC resulting from *E. faecium* bacterial infection (e.g. urinary tract  
 CC infection, bacteraemia, endocarditis, wounds and abdominal/pelvic  
 CC infection), and for screening drugs such as agonists and antagonists. The  
 CC nucleic acid is useful for recombinant production of *Candida albicans* -  
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions  
 CC and vaccines containing the nucleic acid are useful for preventing or  
 CC treating *Enterococcus faecium* infections. The present sequence represents  
 CC one if the disclosed *E. faecium* proteins.  
 XX  
 SQ Sequence 186 AA;  
 XX  
 Query Match 4.9%; Score 7; DB 7; Length 186;  
 Best Local Similarity 100.0%; Pred. No. 60;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 TAKVRR 133  
| | | | |  
Db 43 TAKVRR 49

RESULT 58  
AAU35105  
ID AAU35105 standard; protein; 194 AA.

AC AAU35105;  
XX  
XX  
DT 13-FEB-2002 (first entry)  
XX  
XX  
DE Enterococcus faecalis cellular proliferation protein #392.  
XX  
XX  
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;  
XX antibacterial; drug design.  
XX  
OS Enterococcus faecalis.  
XX  
PN WO200170955-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 21-MAR-2001; 2001WO-US009180.  
XX  
PR 21-MAR-2000; 2000US-0191078P.  
XX  
PR 23-MAY-2000; 2000US-0206848P.  
XX  
PR 26-MAY-2000; 2000US-0207727P.  
XX  
PR 23-OCT-2000; 2000US-0242578P.  
XX  
PR 27-NOV-2000; 2000US-0253625P.  
XX  
PR 22-DEC-2000; 2000US-0257931P.  
XX  
PR 16-FEB-2001; 2001US-0269308P.  
XX  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;  
PI Yamamoto RT, Xu HH;  
XX  
XX  
DR WPI; 2001-611495/70.  
XX  
DR N-PSDB; AAS52964.

PT New polynucleotides for the identification and development of  
XX antibiotics, comprise sequences of antisense nucleic acids.  
XX  
XX  
PS Example 3; SEQ ID NO 10698; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the genes,  
CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
CC useful for the identification of potential new targets for antibiotic  
CC development. The antisense nucleic acids can also be used to identify  
CC proteins used in proliferation, to express these proteins, and to obtain  
CC antibodies capable of binding to the expressed proteins. The proteins can  
CC be used to screen compounds in rational drug discovery programmes. The  
CC antisense nucleic acid sequence is also useful to screen for homologous  
CC nucleic acids which are required for cell proliferation in a wide variety  
CC of organisms. The present sequence represents an essential prokaryotic  
CC cellular proliferation protein. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_ptc\_sequences  
XX  
XX  
SQ Sequence 194 AA;

Query Match 4.9%; Score 7; DB 4; Length 194;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 TAKVRR 133  
| | | | |  
Db 51 TAKVRR 57

RESULT 59  
ABG12973  
ID ABG12973 standard; protein; 220 AA.

AC ABG12973;  
XX  
XX  
DT 13-FEB-2002 (first entry)  
XX  
XX  
DE Novel human diagnostic protein #12964.  
XX  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
XX  
PR 31-MAR-2000; 2000US-00540217.  
XX  
PR 23-AUG-2000; 2000US-00649167.  
XX  
XX  
PA (HYSE-) HISEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
XX  
DR WPI; 2001-639362/73.  
XX  
DR N-PSDB; AAS77160.

PT New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX  
XX  
PS Claim 20; SEQ ID NO 43332; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_ptc\_sequences  
XX  
XX  
SQ Sequence 220 AA;

Query Match 4.9%; Score 7; DB 4; Length 220;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 186 APTLPPA 192

## RESULT 60

AAW65048 ID AAW65048 standard; protein; 307 AA.

AC AAW65048;

DT 28-SEP-1998 (first entry)

DE Tsuga pinoresinol-lariciresinol reductase PLR-Th1.

KW pinoresinol-lariciresinol reductase; dirigent protein; lignan;  
transgenic plant; anticancer; cytostatic; antiviral; virucide;  
antibiotic; antioxidant; antifedant.

OS Tsuga heterophylla.

PN WO9820113-A1.

PD 14-MAY-1998.

PF 07-NOV-1997; 97WO-US020391.

PR 08-NOV-1996; 96US-0030523P.

PR 31-JUL-1997; 97US-0054380P.

PA (UNIV ) UNIV WASHINGTON STATE RES FOUND.

PI Lewis NG, Davin LB, Dinkovakostova AT, Fujita M, Gang DR;

PI Sarkanen S;

DR MPI: 1998-286929/25.

DR N-PSDB; AAW35186.

PT New plant-derived dirigent proteins and pinoresinol-lariciresinol  
reductases - and related nucleic acid, vectors and transformants, used  
for stereospecific production of lignans, useful e.g. as anticancer or  
antiviral agents.

PS Claim 17; Page 134-135; 148pp; English.

CC PLR-Th1 comprises a (+)-pinoresinol/(+)-lariciresinol reductase (P/LR) of  
CC Tsuga heterophylla. P/LR enzymes catalyse the conversion of pinoresinol  
CC to lariciresinol and then to secoisolariciresinol. 2 isoforms (see  
CC AAW65048-49) of T. heterophylla P/LR have been identified. Further P/LR  
CC enzymes have been obtained from Forsythia intermedia (see AAW65038-43)  
CC and Thujia plicata (see AAW65044-47). The isolation of cDNAs encoding P/LR  
CC (see AAV35175-80 and AAV35182-87) and dirigent proteins (see AAV35160-71)  
CC permits the development of an efficient expression system for these  
CC enzymes. Provides useful tools for examining the developmental regulation  
CC of lignan biosynthesis and permits the isolation of related sequences. It  
CC also allows the transformation of a wide range of organisms, including  
CC plants, in order to modify lignan biosynthesis. Optically pure lignans  
CC may have e.g. anticancer, antiviral, antioxidant, antibiotic or  
CC antifedant activity

SQ Sequence 307 AA;

Query Match 4.9%; Score 7; DB 2; Length 307;

Best Local Similarity 100.0%; Pred. No. 95;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAIE 136

Db 135 KVRRAIE 141

## RESULT 61

AAE06561 ID AAE06561 standard; protein; 307 AA.

XX AAE06561;

DT 25-SEP-2001 (first entry)

DE Tsuga heterophylla pinoresinol/lariciresinol reductase, p1r-Tp3 #2.

KW Dirigent protein; pinoresinol/lariciresinol reductase; stereospecificity;  
lignan biosynthetic pathway; secoisolariciresinol.

OS Tsuga heterophylla.

PN WO200149833-A2.

PD 12-JUL-2001.

PF 22-DEC-2000; 2000WO-US035265.

PR 30-DEC-1999; 99US-00475316.

PA (UNIV ) UNIV WASHINGTON STATE RES FOUND.

PA (MIND ) UNIV MINNESOTA.

PI Lewis NG, Davin LB, Dinkova-Kostova AT, Fujita M, Gang DR;

PI Ford JD, Sarkanen S;

DR MPI: 2001-465260/50.

DR N-PSDB; AAD12522.

PT Dirigent and/or pinoresinol/lariciresinol reductase proteins useful for  
producing optically-pure lignans.

PS Claim 24; Page 151-152; 183pp; English.

CC The present invention relates to an isolated dirigent and/or pinoresinol  
CC /lariciresinol reductase protein from a lignan biosynthetic pathway.  
CC Dirigent and/or pinoresinol/lariciresinol reductase protein and the  
CC nucleic acids that encode it may be expressed either in vivo or in vitro  
CC to produce enzymes involved in the biosynthesis of lignans. The 78-kD  
CC dirigent protein confers stereospecificity in 8,8'-linked lignan  
CC formation and binds to and orients conferyl alcohol-derived free  
CC radicals, which then under go stereospecific coupling to form (+)-  
CC pinoresinol. Pinoresinol/lariciresinol reductase catalyses the conversion  
CC of pinoresinol to lariciresinol and then to secoisolariciresinol. The  
CC present sequence is Tsuga heterophylla pinoresinol/lariciresinol  
CC reductase, p1r-Tp3 #2

SQ Sequence 307 AA;

Query Match 4.9%; Score 7; DB 4; Length 307;

Best Local Similarity 100.0%; Pred. No. 95;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAIE 136

Db 135 KVRRAIE 141

## RESULT 62

AAV44811 ID AAV44811 standard; protein; 308 AA.

AC AAV44811;

DT 18-MAY-2000 (first entry)

DE P. taeda dehydrodiconiferyl alcohol benzylic ether reductase.

KW Dehydrodiconiferyl alcohol benzylic ether reductase; lignan content;  
lignan biosynthesis; pulp and paper production; loblolly pine;  
recombinant expression vector; dehydrodiconiferyl alcohol; DDC.

OS Pinus taeda.

XX W0200005350-A1.  
 XX 03-FEB-2000.  
 XX 23-JUL-1999; 99WO-US016746.  
 XX 24-JUL-1998; 98US-0094012P.  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX Lewis NG, Kasahara H, Gang DR, Davin LB;  
 XX WPI; 2000-182684/16.  
 XX N-PSDB; AA250238.  
 XX  
 PT New nucleic acid sequence encoding an enzyme which converts  
 PT dehydrodiconiferyl alcohol to 7-O-4'-isodihydrodiconiferyl alcohol  
 PT useful for producing transgenic plants with modified lignan biosynthesis.  
 XX  
 PS Claim 6; Page 45-46; 55pp; English.  
 CC The present sequence is Pinus taeda dehydrodiconiferyl alcohol (DDC)  
 CC benzyl ether reductase. The corresponding cDNA was cloned into the  
 CC expression plasmid pSBETA, expressed in E. coli resulting in this protein  
 CC which lacks a beta-galactosidase domain. This was purified from a crude  
 CC E. coli extract and resulting protein was assayed for pinoresinol-  
 CC lacticresinol activity and its ability to reduce DDC. The purified P.  
 CC taeda reductase effected the reduction of the benzyl ether bond of DDC  
 CC to convert it to 7-O-4'-(iso) DDC. DDC gene when introduced in a host  
 CC cell permits the development of an efficient expression system for the  
 CC enzyme, provides useful tools for examining the developmental regulation  
 CC of lignan biosynthesis and for isolation of similar proteins. This is  
 CC used to modify lignan levels in plants and food items, altering colour,  
 CC texture, durability and pest-resistance of wood tissue, altering the  
 CC lignan/lignan content of plant species utilised in pulp and paper  
 CC production, and enhancing production of defensive lignans/lignins  
 XX  
 SQ Sequence 308 AA;  
 Query Match 4.9%; Score 7; DB 3; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 130 KVRRAIE 136  
 Db 136 KVRRAIE 142  
 RESULT 63  
 AA65049  
 ID AA65049 standard; protein, 309 AA.  
 XX  
 AC AA65049;  
 XX  
 DT 28-SEP-1998 (first entry)  
 XX  
 DE Tsuga pinoresinol-lariciresinol reductase pLR-Th2.  
 XX  
 KW Pinoresinol-lariciresinol reductase; dirigent protein; lignan;  
 KW transgenic plant; anticancer; cyostatic; antiviral; virucide;  
 KW antibiotic; antioxidant; antifedant.  
 XX  
 OS Tsuga heterophylla.  
 XX  
 OS WO9820113-A1.  
 XX  
 PD 14-MAY-1998.  
 XX  
 PF 07-NOV-1997; 97WO-US020391.  
 XX  
 PF 08-NOV-1996; 96US-0030522P.  
 XX  
 PR 31-JUL-1997; 97US-0054380P.  
 XX

XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX Lewis NG, Davin LB, Dinkovakostova AT, Fujita M, Gang DR;  
 XX Sarkanen S;  
 XX WPI; 1998-286928/25.  
 XX N-PSDB; AAV35187.  
 XX  
 PT New plant-derived dirigent proteins and pinoresinol-lariciresinol  
 PT reductases - and related nucleic acid, vectors and transformants, used  
 PT for stereospecific production of lignans, useful e.g. as anticancer or  
 PT antiviral agents.  
 XX  
 PS Claim 17; Page 138-139; 148pp; English.  
 CC  
 CC pLR-Th2 comprises a (+)-pinoresinol/(+)-lariciresinol reductase (P/LR) of  
 CC Tsuga heterophylla. P/LR enzymes catalyse the conversion of pinoresinol  
 CC to lariciresinol and then to secoisolariciresinol. 2 isoflorms (see  
 CC AA65048-49) of T. heterophylla P/LR have been identified. Further P/LR  
 CC enzymes have been obtained from Forsythia intermedia (see AA65038-43)  
 CC and Thuja plicata (see AA65044-47). The isolation of cDNAs encoding P/LR  
 CC (see AAV35175-80 and AAV35182-87) and dirigent proteins (see AAV35160-71)  
 CC permits the development of an efficient expression system for these  
 CC enzymes, provides useful tools for examining the developmental regulation  
 CC of lignan biosynthesis and permits the isolation of related sequences. It  
 CC also allows the transformation of a wide range of organisms, including  
 CC plants, in order to modify lignan biosynthesis. Optically pure lignans  
 CC may have e.g. anticancer, antiviral, antioxidant, antibiotic or  
 CC antifeedant activity  
 XX  
 SQ Sequence 309 AA;  
 Query Match 4.9%; Score 7; DB 2; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 130 KVRRAIE 136  
 Db 137 KVRRAIE 143  
 RESULT 64  
 AA60562  
 ID AA60562 standard; protein, 309 AA.  
 XX  
 AC AA60562;  
 XX  
 DT 25-SEP-2001 (first entry)  
 XX  
 DE Tsuga heterophylla pinoresinol/lariciresinol reductase, pLR-Tp4 #2.  
 XX  
 KW Dirigent protein; pinoresinol/lariciresinol reductase; stereospecificity;  
 KW lignan biosynthetic pathway; secoisolariciresinol.  
 XX  
 OS Tsuga heterophylla.  
 XX  
 PN WO200149833-A2.  
 XX  
 PD 12-JUL-2001.  
 XX  
 PF 22-DEC-2000; 2000WO-US035265.  
 XX  
 PF 30-DEC-1999; 99US-00475316.  
 XX  
 PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX  
 PA (MINU ) UNIV MINNESOTA.  
 XX  
 PI Lewis NG, Davin LB, Dinkova-Kostova AT, Fujita M, Gang DR;  
 XX Ford UD, Sarkanen S;  
 XX WPI; 2001-465260/50.  
 XX N-PSDB; AAD12523.  
 DR

The present invention relates to an isolated dirigent and/or pinoreosinol  
 /lariciresinol reductase protein from a lignan biosynthetic pathway.  
 Dirigent and/or pinoreosinol/lariciresinol reductase protein and the  
 nucleic acids that encode it may be expressed either in vivo or in vitro  
 to produce enzymes involved in the biosyntheses of lignans. The 78-kD  
 dirigent protein confers stereospecificity in 8,8'-linked lignan  
 formation and binds to and orients coniferyl alcohol-derived free  
 radicals, which then under go stereospecific coupling to form (+)-  
 pinoreosinol. Pinoreosinol/lariciresinol reductase catalyses the conversion  
 of pinoreosinol to lariciresinol and then to secoisolariciresinol. The  
 present sequence is Tsuga heterophylla pinoreosinol/lariciresinol  
 reductase, pLR-Tp4 #2

Sequence 309 AA;  
 Query Match 4.9%; Score 7; DB 4; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

130 KVRRATE 136  
 |||||  
 137 KVRRATE 143

RESULT 65  
 AAM65044  
 ID AAM65044 standard; protein; 313 AA.  
 XX  
 AC AAM65044;  
 XX  
 DT 28-SEP-1998 (first entry)  
 XX  
 DE Thuja pinoreosinol-lariciresinol reductase pLR-Tp1.  
 XX  
 KW Pinoreosinol-lariciresinol reductase; dirigent protein; lignan;  
 KW transgenic plant; anticancer; cytostatic; antiviral; virucide;  
 KW antibiotic; antioxidant; antifeedant.  
 XX  
 OS Thuja plicata.  
 OS  
 XX WO9820113-A1.  
 XX  
 PD 14-MAY-1998.  
 PD  
 XX 07-NOV-1997; 97MO-US020391.  
 PF  
 XX 08-NOV-1996; 96US-00305222.  
 PR  
 XX 31-JUL-1997; 97US-0054380P.  
 RR  
 XX (UNIV ) UNIV WASHINGTON STATE RES FOUND.  
 PA  
 XX Lewis NG, Davin LB, Dinkovakostova AT, Fujita M, Gang DR;  
 PI Saranen S;  
 PI  
 XX WP1, 1998-286929/25.  
 DR  
 XX N-PSDB; AAV35182.  
 DR  
 XX New plant-derived dirigent proteins and pinoreosinol-lariciresinol  
 PT reductases - and related nucleic acid, vectors and transformants, used  
 PT for stereospecific production of lignans, useful e.g. as anticancer or  
 PT antiviral agents.  
 PT  
 XX  
 XX Claim 17, Page 122-123; 14BPP; English.

pLR-Tp1 comprises a (-)-pinoreosinol/(-)-lariciresinol reductase (P/LR) of  
 Thuja plicata. P/LR enzymes catalyse the conversion of pinoreosinol to  
 lariciresinol and then to secoisolariciresinol. 4 isoforms (see AAM65044-  
 CC

CC 47) of *T. plicata* P/LR have been identified; the other 3 isoforms are  
CC (+)/(+) P/LR. Further P/LR enzymes have been obtained from *Forcythia*  
CC *intermedia* (see AAM65038-43) and *Tsuga heterophylla* (see AAM65048-49).  
CC The isolation of cDNAs encoding P/LR (see AAV35175-80 and AAV35182-87)  
CC and divergent proteins (see AAV35160-71) permits the development of an  
CC efficient expression system for these enzymes, provides useful tools for  
CC examining the developmental regulation of lignan biosynthesis and permits  
CC the isolation of related sequences. It also allows the transformation of  
CC a wide range of organisms, including plants, in order to modify lignan  
CC biosynthesis. Optically pure lignans may have e.g. anticancer, antiviral,  
CC antioxidant, antibiotic or antifeedant activity  
CC  
CC  
SQ Sequence 313 AA;  
Query Match 4.9%; Score 7; DB 2; Length 313;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 130 KYRRATE 136  
|||  
|||  
Db 140 KYRRATE 146  
RESULT 66  
AAAE06557  
ID AAE06557 standard; protein; 313 AA.  
XX  
AC AAE06557;  
XX  
DT 25-SEP-2001 (first entry)  
XX  
DE *Thuja plicata* pinoresinol/lariciresinol reductase, p1r-Tp1.  
XX  
KW Divergent protein; pinoresinol/lariciresinol reductase; stereospecificity;  
KM lignan biosynthetic pathway; secoisolariciresinol.; western red cedar.  
OS  
SS *Thuja plicata*.  
XX  
PN MO200149833-A2.  
XX  
PD 12-JUL-2001.  
XX  
PF 22-DEC-2000; 2000MO-US035265.  
XX  
PR 30-DEC-1999; 99US-00475316.  
XX  
PA (UNIM) UNIV WASHINGTON STATE RES FOUND.  
PA (MINU) UNIV MINNESOTA.  
XX  
PI Lewis NG, Davin LB, Dinkova-Kostova AT, Fujita M, Gang DR,  
PI Ford JD, Sarkanen S;  
XX  
XX *WPI*; 2001-465260/50.  
DR *N*-PSDB; AAD12518.  
XX  
PT Divergent and/or pinoresinol/lariciresinol reductase proteins useful for  
PT producing optically-pure lignans.  
XX  
PS Claim 27; Page 139-140; 183pp; English.  
XX  
XX The present invention relates to an isolated divergent and/or pinoresinol  
XX /lariciresinol reductase protein from a lignan biosynthetic pathway.  
XX Divergent and/or pinoresinol/lariciresinol reductase protein and the  
XX nucleic acids that encode it may be expressed either *in vivo* or *in vitro*  
XX to produce enzymes involved in the biosynthesis of lignans. The 78-kD  
XX divergent protein confers stereospecificity in 8,8'-linked lignan  
XX formation and binds to and orients coniferyl alcohol-derived free  
XX radicals, which then under go stereospecific coupling to form (+)-  
XX pinoresinol. Pinoresinol/lariciresinol reductase catalyses the conversion  
XX of pinoresinol to lariciresinol and then to secoisolariciresinol. The  
XX present sequence is *Thuja plicata* pinoresinol/lariciresinol reductase,  
XX p1r-Tp1



DE Schiandra chinensis pinoresinol/lariciresinol reductase protein #7.  
 XX Dirigent protein; pinoresinol/lariciresinol reductase; stereospecificity;  
 KM lignan biosynthetic pathway; secoisolariciresinol.  
 XX Schiandra chinensis.  
 XX MO200149833-A2.  
 XX 12-UTL-2001.  
 XX 22-DEC-2000; 2000WO-US035265.  
 XX 30-DEC-1999; 99US-00475316.  
 XX (UNIV ) UNIV WASHINGTON STATE RES FOUND.  
 PA (MINU ) UNIV MINNESOTA.  
 XX Lewis NG, Davin LB, Dinkova-Kostova AT, Fujita M, Gang DR;  
 PI Ford JD, Sarkanen S;  
 XX MPI; 2001-465260/50.  
 DR N-PSDB; AAD12561.  
 XX PT Dirigent and/or pinoresinol/lariciresinol reductase proteins useful for  
 PT producing optically-pure lignans.  
 XX Claim 33; Page 180-181; 183pp; English.  
 XX The present invention relates to an isolated dirigent and/or pinoresinol  
 CC /lariciresinol reductase protein from a lignan biosynthetic pathway.  
 CC Dirigent and/or pinoresinol/lariciresinol reductase protein and the  
 CC nucleic acids that encode it may be expressed either in vivo or in vitro  
 CC to produce enzymes involved in the biosynthesis of lignans. The 78-kD  
 CC dirigent protein confers stereospecificity in 8,8'-linked lignan  
 CC formation and binds to and orients coniferyl alcohol-derived free  
 CC radicals, which then under go stereospecific coupling to form (+)-  
 CC pinoresinol. Pinoresinol/lariciresinol reductase catalyzes the conversion  
 CC of pinoresinol to lariciresinol and then to secoisolariciresinol. The  
 CC present sequence is Schiandra chinensis pinoresinol/lariciresinol  
 CC reductase  
 XX  
 SO Sequence 314 AA;  
 Query Match 4.9%; Score 7; DB 4; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 97;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 130 KVRRAIE 136  
 Db 142 KVRRAIE 148  
 RESULT 70  
 AAY32075  
 ID AAY32075 standard; protein; 572 AA.  
 XX  
 AC AAY32075;  
 XX  
 DT 17-JUN-2000 (first entry)  
 XX  
 DE Rapeseed raffinose synthase.  
 XX  
 KW Raffinose synthase; rapeseed; transgenic plant.  
 XX  
 OS Brassica napus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 129  
 FT Misc-difference 132 /note= "encoded by GGY"  
 FT Misc-difference 132 /note= "encoded by GGY"  
 FT Misc-difference 143

FT /note= "encoded by CCS"  
 FT Misc-difference 144  
 FT /note= "encoded by TCR"  
 FT Misc-difference 148  
 FT /note= "encoded by CGR"  
 XX  
 XX EP953643-A2.  
 XX 03-NOV-1999.  
 XX 27-APR-1999; 99EP-00107430.  
 XX 30-APR-1998; 98JP-00120550.  
 XX 30-APR-1998; 98JP-00120551.  
 XX 04-DEC-1998; 98JP-00345590.  
 XX 10-DEC-1998; 98JP-00351246.  
 XX (SUMO ) SUMITOMO CHEM CO LTD.  
 XX Watanabe E, Oeda K;  
 PI MPI; 1999-593144/51.  
 DR N-PSDB; AAZ20210.  
 XX PT New sense and antisense genes, useful for altering the level of raffinose  
 PT in food plants.  
 XX Claim 27; Page 36-38; 55pp; English.  
 XX This sequence represents rapeseed raffinose synthase, a protein that can  
 CC bind a D-galactosyl group through an alpha(1-6) bond to the hydroxy group  
 CC attached to the carbon atom at the 6-position of the D-glucose residue in  
 CC an sucrose molecule to form raffinose. cDNA (see AAZ20210) encoding the  
 CC enzyme was isolated from rapeseed cv. Westar leaf cDNA by PCR. Probes or  
 CC primers generated from plant raffinose synthase genes (see AAZ20207-10)  
 CC may be used to obtain other raffinose synthase genes by labeled detection  
 CC or amplification (claimed). These genes may be used to control the levels  
 CC of raffinose produced in plants. Antisense genes can be used to knock out  
 CC existing gene activity, and sense genes to increase the level of gene  
 CC activity. The resulting transgenic plants may be used as a food source to  
 CC alter the growing conditions for gut enterobacteria, providing general  
 CC health advantages  
 XX  
 SO Sequence 572 AA;  
 Query Match 4.9%; Score 7; DB 2; Length 572;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 GAPTLP 8  
 Db 137 GAPTLP 143  
 RESULT 71  
 ADC96107  
 ID ADC96107 standard; protein; 580 AA.  
 XX  
 AC ADC96107;  
 XX  
 DT 01-JUN-2004 (first entry)  
 XX  
 DE E. faecium protein sequence SEQ ID 5734.  
 XX  
 KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;  
 KM abdominal-pelvic infection.  
 XX  
 OS Enterococcus faecium.  
 XX  
 FH US6583275-B1.  
 XX 24-JUN-2003.  
 PD

PF 30-JUN-1998; 98US-00107532.  
 XX  
 PR 02-JUL-1997; 97US-0051571P.  
 PR 14-MAY-1998; 98US-0085598P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Doucette-Stamm LA, Bush D;  
 XX  
 DR WPI; 2003-799836/75.  
 DR N-PSDB; ADC92453.  
 XX  
 PT New isolated nucleic acid derived from *Enterococcus faecium* encoding an  
 PT *Enterococcus faecium* polypeptide useful for detection, prevention and  
 PT treatment of a pathological condition resulting from a bacterial  
 PT infection.  
 XX  
 PS Example 1; SEQ ID NO 5734; 243pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid derived from  
 CC *Enterococcus faecium* encoding an *Enterococcus faecium* polypeptide having  
 CC one of 10 fully defined sequences given in the (or comprising 40  
 CC sequential nucleotides chosen from any of the nucleic acids, its  
 CC complement or sequences hybridizing to it). Also included are a  
 CC recombinant vector comprising the nucleic acid operably linked to a  
 CC transcription regulatory element, a cell comprising the vector and a  
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are  
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.  
 CC The nucleic acids is useful for diagnosing pathological conditions  
 CC resulting from *E. faecium* bacterial infection (e.g. urinary tract  
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic  
 CC infection) and for screening drugs such as agonists and antagonists. The  
 CC nucleic acid is useful for recombinant production of *Candida albicans* -  
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions  
 CC and vaccines containing the nucleic acid are useful for preventing or  
 CC treating *Enterococcus faecium* infections. The present sequence represents  
 CC one if the disclosed *E. faecium* proteins.  
 XX  
 SQ Sequence 580 AA;  
 XX  
 QY  
 DB 111 NRIAKET 117  
 227 NRIAKET 233  
 XX  
 RESULT 72  
 ID AAY32074 standard; protein; 777 AA.  
 AC AAY32074;  
 XX  
 DT 17-JAN-2000 (first entry)  
 DE Mustard raffinose synthase.  
 XX  
 KM Raffinose synthase; mustard; transgenic plant.  
 KM  
 XX  
 OS Brassica juncea.  
 XX  
 FH Key Location/Qualifiers  
 FT MISC-difference 210 /note= "encoded by ACR"  
 FT  
 XX  
 PN BP953643-A2.  
 XX  
 PD 03-NOV-1999.  
 XX  
 PF 27-APR-1999; 99EP-00107430.  
 XX

PR 30-APR-1998; 98EP-00120550.  
 PR 30-APR-1998; 98EP-00120551.  
 PR 04-DEC-1998; 98EP-00345580.  
 PR 10-DEC-1998; 98EP-00351246.  
 XX  
 PA (SUMO) SUMITOMO CHEM CO LTD.  
 XX  
 PI Watanabe E, Oeda K;  
 XX  
 DR WPI; 1999-593144/51.  
 DR N-PSDB; AA220209.  
 XX  
 PT New sense and antisense genes, useful for altering the level of raffinose  
 PT in food plants.  
 XX  
 PS Claim 26; Page 29-31; 55pp; English.  
 XX  
 CC This sequence represents mustard raffinose synthase, a protein that can  
 CC bind a D-galactosyl group through an alpha(1-6) bond to the hydroxy group  
 CC attached to the carbon atom at the 6-position of the D-glucose residue in  
 CC a sucrose molecule to form raffinose. cDNA (see AA220209) encoding the  
 CC enzyme was isolated from mustard (*Brassica juncea*) leaf cDNA by PCR.  
 CC Probes or primers generated from plant raffinose synthase genes (see  
 CC AA220207-10) may be used to obtain other raffinose synthase genes by  
 CC labeled detection or amplification (claimed). These genes may be used to  
 CC control the levels of raffinose produced in plants. Antisense genes can  
 CC be used to knock out existing gene activity and sense genes to increase  
 CC the level of gene activity. The resulting transgenic plants may be used  
 CC as a food source to alter the growing conditions for gut enterobacteria,  
 CC providing general health advantages  
 XX  
 SQ Sequence 777 AA;  
 XX  
 QY  
 DB 2 GAPTLP 8  
 342 GAPTLP 348  
 XX  
 RESULT 73  
 ID AB014746 standard; protein; 823 AA.  
 AC AB014746;  
 XX  
 DT 25-AUG-2003 (first entry)  
 DE Novel human protein #119.  
 XX  
 KM Human; NOV; gene therapy; endocrine related disease; diabetes;  
 KM metabolism-related disease; obesity; central nervous system disorder;  
 KM Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;  
 KM schizophrenia; depression; autoimmune disorder; inflammatory disorder;  
 KM psoriasis; allergy; lupus erythematosus; asthma; cancer;  
 KM inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;  
 KM colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;  
 KM prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;  
 KM lung disease; emphysema; obstructive pulmonary disease; haemophilia;  
 KM stroke; infection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003023002-A2.  
 XX  
 PD 20-MAR-2003.  
 XX  
 PF 09-SEP-2002; 2002WO-US028539.  
 XX  
 PR 07-SEP-2001; 2001US-0318120P.  
 PR 07-SEP-2001; 2001US-0318130P.  
 PR



PR 10-SEP-2001; 2001US-0318430P.  
PR 17-SEP-2001; 2001US-0322636P.  
PR 17-SEP-2001; 2001US-0322781P.  
PR 17-SEP-2001; 2001US-0322816P.  
PR 17-SEP-2001; 2001US-0322817P.  
PR 17-SEP-2001; 2001US-0322817P.  
PR 19-SEP-2001; 2001US-0323519P.  
PR 20-SEP-2001; 2001US-0323631P.  
PR 20-SEP-2001; 2001US-0323631P.  
PR 25-SEP-2001; 2001US-0324969P.  
PR 25-SEP-2001; 2001US-0325091P.  
PR 26-SEP-2001; 2001US-0324969P.  
PR 17-APR-2002; 2002US-0373212P.  
PR 06-SEP-2002; 2002US-00236177.  
(CUBA-) CUBAGEN CORP.

XX Spytek KA, Patutajan M, Gorman L, Li L, Anderson DW, Zhong M;  
PI Gerlach VL, Vernet CM, Ellerman K, Berghs C, Rothenberg ME, Guo X;  
PI Shinkens RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;  
PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru W, Alsobrook JP;  
PI Lepley DM, Edinger SR, Burgess CB;  
XX WPI: 2003-313242/30.  
DR N-PSDB; ACD19439.

PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)  
PT and polynucleotides, useful in gene therapy, e.g. for treating or  
PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,  
PT stroke or infections.

PS Claim 1; Page 334; 586pp; English.

XX The invention describes a new isolated polypeptide (NOVX). The NOVX  
CC polypeptide, nucleic acid and antibody are useful as therapeutics,  
CC particularly in the manufacture of a medicament for treating a syndrome  
CC associated with a human disease, which includes a pathology associated  
CC with NOVX polypeptide. The DNA encoding the protein is useful in gene  
CC therapy for treating the disease or condition. In particular, the NOVX  
CC polypeptide or polynucleotide is useful for treating endocrine/  
CC metabolism-related diseases (e.g. obesity or diabetes), central nervous  
CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,  
CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune  
CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,  
CC asthma, inflammatory bowel disease, rheumatoid arthritis or  
CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,  
CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver  
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),  
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).  
CC These are also useful in developing powerful assay system for functional  
CC analysis of various human disorders, as well as in diagnostic  
CC applications, and for monitoring the effects of drugs during clinical  
CC trials. This is the amino acid sequence of a novel human NOV protein  
XX

SO Sequence 823 AA;

Query Match 4.9%; Score 7; DB 6; Length 823;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 KIRAKETN 118  
Db 547 KIRAKETN 553

RESULT 74  
ID ABJ25599  
XX ABJ25599 standard; protein; 838 AA.  
AC  
XX ABJ25599;  
DT 16-APR-2003 (first entry)  
XX Aspergillus fumigatus essential gene protein #257.

XX Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;  
KW Cancer; contamination; biofilm; antibody; immune response.

XX Aspergillus fumigatus.

OS WO200286090-A2.

FN 31-OCT-2002.

PD 23-APR-2002; 2002WO-US013142.

PF 23-APR-2001; 2001US-0285697P.

PR 27-APR-2001; 2001US-0287066P.

PR 05-JUN-2001; 2001US-0295890P.

PR 09-JUL-2001; 2001US-0303899P.

PR 31-AUG-2001; 2001US-0316362P.

PA (ELIT-) ELITRA PHARM INC.

PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

XX WPI: 2003-093124/08.

DR New purified or isolated nucleic acids of essential genes of Aspergillus

PT fumigatus, useful for treating or preventing infections by A. fumigatus,

PT or for treating a non-infectious disease in a subject e.g. cancer.

PT Disclosure; Page: 175pp; English.

XX The invention relates to novel purified or isolated nucleic acids of  
CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of  
CC the invention are used to treat or prevent infections by a pathogenic  
CC organism such as A. fumigatus, to treat a non-infectious disease in a  
CC subject (e.g. cancer), to prevent or contain contamination of an object  
CC by A. fumigatus, or to prevent or inhibit formation on a surface of a  
CC biofilm comprising A. fumigatus. The polynucleotides are useful for  
CC expressing recombinant protein for characterization, screening or  
CC therapeutic use, as markers for host tissues in which the pathogenic  
CC organisms invade or reside, for comparing with the DNA sequence of A.  
CC fumigatus to identify duplicated genes or paralogues having the same or  
CC similar biochemical activity and/or function, for comparing with DNA  
CC sequences of other related or distant pathogenic organisms to identify  
CC potential orthologous essential or virulence genes, for selecting and  
CC making oligomers for attachment to a nucleic acid array for examination  
CC of expression patterns, for raising anti-protein antibodies, as an  
CC antigen to raise anti-DNA antibodies or to elicit another immune  
CC response, and for identifying polynucleotides encoding the other protein  
CC with which binding occurs or to identify inhibitors of the binding  
CC interaction. The polypeptides may be used to raise antibodies or to  
CC elicit immune response, as a reagent in assays designed to quantitatively  
CC determine levels of the protein in biological fluids, as a marker for  
CC host tissues in which pathogenic organism invade or reside, and to  
CC isolate correlative receptors or ligands in the case of virulence  
CC factors. This sequence represents a protein of one of the essential genes  
CC of Aspergillus fumigatus of the invention  
XX

SO Sequence 838 AA;

Query Match 4.9%; Score 7; DB 6; Length 838;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 EELTIGE 100  
Db 11 EELTIGE 17

RESULT 75  
ID ABB61778  
XX ABB61778 standard; protein; 1205 AA.  
AC  
XX ABB61778;

```

XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 12126.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX W0200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (BEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX N-PSDB; ABL05881.
XX
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
XX Disclosure; SEQ ID NO 12126; 21pp + Sequence listing; English.
XX
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX ABBS7072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at tcp.wipo.int/pub/published_pct_sequences
XX
XX
XX Sequence 1205 AA;
XX
XX
XX Query Match 4.9%; Score 7; DB 4; Length 1205;
XX Best Local Similarity 100.0%; Pred. No. 3.4e+02;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 3 APTLPBA 9
XX
XX 418 APTLPBA 424

```

Search completed: August 11, 2004, 14:23:52  
 Job time : 57 secs

Tue Aug 17 05:54:49 2004

us-09-690-825-34.oligo.rapb

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 14:25:13 : Search time 46 Seconds  
(Without alignments)  
969.079 Million cell updates/sec

Title: US-09-690-825-34  
Perfect score: 142  
Sequence: 1 MGAPLTPAWQPLKDHRI...EPEETAKVRAIEQIAMD 142

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 1292805 seqs, 313927144 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database:

Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PTCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PTCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142	100.0	142	US-10-138-618-34	Sequence 34, Appl
2	142	100.0	142	US-10-665-975-2	Sequence 2, Appl
3	128	90.1	142	US-10-408-765A-2210	Sequence 2210, Ap
4	98	69.0	98	US-09-965-967-23	Sequence 23, Appl
5	20	14.1	20	US-10-138-618-4	Sequence 4, Appl
6	17	12.0	17	US-10-138-618-3	Sequence 1, Appl
7	17	12.0	17	US-10-291-607-12	Sequence 12, Appl
8	9	6.3	9	US-10-291-607-12	Sequence 12, Appl
9	9	6.3	9	US-10-291-607-12	Sequence 12, Appl
10	7	4.9	68	US-10-424-599-148276	Sequence 148276,
11	7	4.9	73	US-09-864-761-38999	Sequence 38999, A
12	7	4.9	99	US-10-424-599-178993	Sequence 178993,
13	7	4.9	101	US-09-933-767-450	Sequence 450, App
14	7	4.9	101	US-10-004-860-450	Sequence 450, App
15	7	4.9	101	US-10-023-282-450	Sequence 450, App

16	7	4.9	126	US-10-424-599-209010	Sequence 209010,
17	7	4.9	176	US-10-437-963-116190	Sequence 116190,
18	7	4.9	179	US-10-282-122A-60941	Sequence 60941, A
19	7	4.9	182	US-10-282-122A-57125	Sequence 57125, A
20	7	4.9	182	US-10-282-122A-57125	Sequence 57125, A
21	7	4.9	185	US-10-424-599-245457	Sequence 245457,
22	7	4.9	194	US-09-815-242-10698	Sequence 10698, A
23	7	4.9	199	US-10-335-977-7963	Sequence 7963, Ap
24	7	4.9	203	US-10-335-977-7964	Sequence 7964, Ap
25	7	4.9	259	US-10-424-599-243648	Sequence 243648,
26	7	4.9	311	US-10-424-599-147883	Sequence 147883,
27	7	4.9	311	US-10-425-114-50128	Sequence 50128, A
28	7	4.9	382	US-10-369-493-13992	Sequence 13992, A
29	7	4.9	388	US-10-424-599-154013	Sequence 154013,
30	7	4.9	415	US-10-424-599-248527	Sequence 248527,
31	7	4.9	428	US-10-421-654-28	Sequence 28,
32	7	4.9	428	US-10-421-654-28	Sequence 28,
33	7	4.9	475	US-10-369-493-10000	Sequence 10000, A
34	7	4.9	610	US-10-437-963-192686	Sequence 192686,
35	7	4.9	744	US-10-369-493-21957	Sequence 21957, A
36	7	4.9	751	US-10-437-963-201149	Sequence 201149,
37	7	4.9	802	US-10-437-963-180519	Sequence 180519,
38	7	4.9	838	US-10-128-714-3257	Sequence 3257, Ap
39	7	4.9	2285	US-09-932-183A-2	Sequence 2,
40	6	4.2	12	US-10-083-768-130	Sequence 130, App
41	6	4.2	30	US-09-939-960-380	Sequence 380, App
42	6	4.2	36	US-09-989-919-109	Sequence 109, App
43	6	4.2	42	US-10-424-599-162708	Sequence 162708,
44	6	4.2	52	US-10-424-599-276718	Sequence 276718,
45	6	4.2	52	US-10-425-114-68762	Sequence 68762, A
46	6	4.2	55	US-09-860-352A-11	Sequence 11, Appl
47	6	4.2	55	US-10-424-599-152902	Sequence 152902,
48	6	4.2	55	US-10-410-764-98	Sequence 98, Appl
49	6	4.2	57	US-10-424-599-158384	Sequence 158384,
50	6	4.2	57	US-10-424-599-352199	Sequence 352199,
51	6	4.2	58	US-10-125-258-62	Sequence 62, Appl
52	6	4.2	58	US-10-424-599-167067	Sequence 167067,
53	6	4.2	58	US-10-424-599-271511	Sequence 271511,
54	6	4.2	58	US-10-437-963-183020	Sequence 183020,
55	6	4.2	62	US-10-424-599-194266	Sequence 194266,
56	6	4.2	63	US-10-424-599-237412	Sequence 237412,
57	6	4.2	63	US-10-125-258-62	Sequence 62, Appl
58	6	4.2	64	US-10-425-114-54599	Sequence 54599, A
59	6	4.2	65	US-09-764-881-122	Sequence 122, App
60	6	4.2	65	US-09-764-875-1113	Sequence 1113, Ap
61	6	4.2	65	US-10-424-599-281627	Sequence 281627,
62	6	4.2	65	US-09-764-881-122	Sequence 122, App
63	6	4.2	65	US-10-242-747-122	Sequence 122, App
64	6	4.2	65	US-10-125-258-62	Sequence 62, App
65	6	4.2	66	US-10-424-599-199019	Sequence 199019,
66	6	4.2	67	US-10-424-599-167358	Sequence 167358,
67	6	4.2	69	US-10-424-599-153000	Sequence 153000,
68	6	4.2	74	US-10-437-963-141151	Sequence 141151,
69	6	4.2	75	US-10-282-122A-45175	Sequence 45175, A
70	6	4.2	76	US-10-437-963-182219	Sequence 182219,
71	6	4.2	77	US-10-424-599-144105	Sequence 144105,
72	6	4.2	82	US-10-424-599-192922	Sequence 192922,
73	6	4.2	84	US-10-424-599-152919	Sequence 152919,
74	6	4.2	88	US-09-764-891-3922	Sequence 3922, Ap
75	6	4.2	87	US-10-424-599-283512	Sequence 283512,
76	6	4.2	86	US-10-424-599-212617	Sequence 212617,
77	6	4.2	85	US-10-424-599-216095	Sequence 216095,
78	6	4.2	92	US-10-424-599-151600	Sequence 151600,
79	6	4.2	95	US-10-424-599-1358	Sequence 1358, Ap
80	6	4.2	99	US-09-864-408A-1358	Sequence 147395,
81	6	4.2	99	US-10-437-963-17395	Sequence 17395,
82	6	4.2	100	US-09-841-132-563	Sequence 563, App
83	6	4.2	102	US-10-424-599-199829	Sequence 199829,
84	6	4.2	102	US-10-437-963-162484	Sequence 162484,
85	6	4.2	103	US-10-437-963-16812	Sequence 16812,
86	6	4.2	104	US-10-437-963-16812	Sequence 16812,
87	6	4.2	105	US-10-437-963-174606	Sequence 174606,
88	6	4.2	105	US-10-437-963-174606	Sequence 174606,

89 6 4.2 106 14 US-10-083-357-986 Sequence 986, App  
90 6 4.2 107 12 US-10-424-559-283306 Sequence 283306,  
91 6 4.2 108 12 US-10-282-122A-76441 Sequence 76441, A  
92 6 4.2 111 16 US-10-437-963-119889 Sequence 119889,  
93 6 4.2 114 12 US-10-424-559-249704 Sequence 249704,  
94 6 4.2 116 16 US-10-437-963-155767 Sequence 155767,  
95 6 4.2 118 16 US-10-437-963-121763 Sequence 121763,  
96 6 4.2 120 16 US-10-437-963-124901 Sequence 124901,  
97 6 4.2 121 12 US-10-424-559-163067 Sequence 163067,  
98 6 4.2 123 12 US-10-424-559-186413 Sequence 186413,  
99 6 4.2 124 16 US-10-437-963-181095 Sequence 181095,  
100 6 4.2 127 12 US-10-335-977-7217 Sequence 7217, Ap

## ALIGNMENTS

RESULT 1  
US-10-138-618-34  
; Sequence 34, Application US/10138618  
; Publication No. US20030100525A1  
; GENERAL INFORMATION:  
; APPLICANT: Altieri, Dario C.  
; TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS  
; CELLULAR APOPTOSIS, AND ITS MODULATION  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN, LEWIS & BOCKTUS LLP  
; STREET: 1800 M Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036-5869  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/138,618  
; FILING DATE: 06-May-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/975,080  
; FILING DATE: 20-NOV-1997  
; APPLICATION NUMBER: US 60/031,435  
; FILING DATE: 20-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adler, Reid G.  
; REGISTRATION NUMBER: 30,988  
; REFERENCE/DOCKET NUMBER: 044574-5022-01-WO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-467-7000  
; TELEFAX: 202-467-7176  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 142 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:  
US-10-138-618-34  
Query Match 100.0%; Score 142; DB 14; Length 142;  
Best Local Similarity 100.0%; Pred. No. 2.1e-133;  
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGAFTLPAMQPFLEKDHRISTFKMPLFEGCACTPERMAEAGFTICPTNEPDLAQCFPC 60  
DB 1 MGAFTLPAMQPFLEKDHRISTFKMPLFEGCACTPERMAEAGFTICPTNEPDLAQCFPC 60  
QY 61 FKELEGMEPDDDPFIEEHKXSSGCAFLSVKQFEEITLGEFLKIDREKAKNKIAKETNNK 120

DB 61 FKELEGMEPDDDPFIEEHKXSSGCAFLSVKQFEEITLGEFLKIDREKAKNKIAKETNNK 120  
QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
DB 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 2  
US-10-665-975-2  
; Sequence 2, Application US/10665975  
; Publication No. US20040138119A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; APPLICANT: Tamm, Ingo  
; TITLE OF INVENTION: USE OF HEPATITIS VIRUS B X-INTERACTING  
; PROTEIN (HBXIP) IN MODULATION OF APOPTOSIS  
; FILE REFERENCE: BURHAM.005A  
; CURRENT APPLICATION NUMBER: US/10/665,975  
; CURRENT FILING DATE: 2003-09-18  
; PRIOR APPLICATION NUMBER: 60/412,109  
; PRIOR FILING DATE: 2002-09-18  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-665-975-2

Query Match 100.0%; Score 142; DB 16; Length 142;  
Best Local Similarity 100.0%; Pred. No. 2.1e-133;  
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAFTLPAMQPFLEKDHRISTFKMPLFEGCACTPERMAEAGFTICPTNEPDLAQCFPC 60  
DB 1 MGAFTLPAMQPFLEKDHRISTFKMPLFEGCACTPERMAEAGFTICPTNEPDLAQCFPC 60  
QY 61 FKELEGMEPDDDPFIEEHKXSSGCAFLSVKQFEEITLGEFLKIDREKAKNKIAKETNNK 120  
DB 61 FKELEGMEPDDDPFIEEHKXSSGCAFLSVKQFEEITLGEFLKIDREKAKNKIAKETNNK 120  
QY 121 KKEFEETAKKVRRAIEQLAAMD 142  
DB 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 3  
US-10-408-765A-2210  
; Sequence 2210, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2210  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-2210

Query Match 90.1%; Score 128; DB 16; Length 142;  
Best Local Similarity 100.0%; Pred. No. 1.9e-119;



TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 3  
US-10-138-618-3

Query Match 12.0%; Score 17; DB 14; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.4e+09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPAPMOPFLKDRI 19  
Db 1 APTLPAPMOPFLKDRI 17

RESULT 7  
US-10-291-607-1  
; Sequence 1, Application US/10291607  
; Publication No. US20030143232A1  
; GENERAL INFORMATION:  
; APPLICANT: Altieri, Dario C.  
; TITLE OF INVENTION: Methods for Selectively Modulating Survivin Apoptosis  
; FILE REFERENCE: 44574-5063-US  
; CURRENT APPLICATION NUMBER: US/10/291,607  
; PRIOR FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US/09/515,514  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (7)  
; OTHER INFORMATION: PHOSPHORYLATION  
US-10-291-607-1

Query Match 12.0%; Score 17; DB 14; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.4e+09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LEGCACTPERMAAGFI 44  
Db 1 LEGCACTPERMAAGFI 17

RESULT 8  
US-10-291-607-12  
; Sequence 12, Application US/10291607  
; Publication No. US20030143232A1  
; GENERAL INFORMATION:  
; APPLICANT: Altieri, Dario C.  
; TITLE OF INVENTION: Methods for Selectively Modulating Survivin Apoptosis  
; FILE REFERENCE: 44574-5063-US  
; CURRENT APPLICATION NUMBER: US/10/291,607  
; PRIOR FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US/09/515,514  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Survivin phosphorylation motif  
US-10-291-607-12

Query Match 6.3%; Score 9; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.2e+06;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 CACTPERMA 39  
Db 1 CACTPERMA 9

RESULT 9  
US-10-291-607-13  
; Sequence 13, Application US/10291607  
; Publication No. US20030143232A1  
; GENERAL INFORMATION:  
; APPLICANT: Altieri, Dario C.  
; TITLE OF INVENTION: Methods for Selectively Modulating Survivin Apoptosis  
; FILE REFERENCE: 44574-5063-US  
; CURRENT APPLICATION NUMBER: US/10/291,607  
; PRIOR FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US/09/515,514  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: Survivin phosphorylation motif  
US-10-291-607-13

Query Match 6.3%; Score 9; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.2e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 CACTPERMA 39  
Db 1 CACTPERMA 9

RESULT 10  
US-10-424-599-148276  
; Sequence 148276, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongsuei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; PRIOR FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 148276  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1) (68)  
; OTHER INFORMATION: unsure at all Xaa locations  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_104915C.1.pcp  
US-10-424-599-148276

Query Match 4.9%; Score 7; DB 12; Length 68;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAIE 136  
Db 14 KVRRAIE 20

RESULT 11  
US-09-864-761-38999  
Sequence 38999, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Hanzel, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Acoomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 38999  
LENGTH: 73  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO ACO05840.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.7  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9  
OTHER INFORMATION: SWISSPROT HIT: P47825, EVALUATE 1.10e+00  
OTHER INFORMATION: EST\_HUMAN HIT: AM361877.1, EVALUATE 4.00e-30  
US-09-864-761-38999

Query Match

4.9%; Score 7; DB 9; Length 73;

Best Local Similarity 100.0%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 111 NRIAKET 117  
DB 1 NRIAKET 7  
RESULT 12  
US-10-424-599-178983  
Sequence 178983, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 178983  
LENGTH: 99  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_132638C.1.pep  
US-10-424-599-178983  
Query Match  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 97 TLGEFLK 103  
DB 17 TLGEFLK 23  
RESULT 13  
US-09-933-767-450  
Sequence 450, Application US/09933767  
Publication No. US20030181632A1  
GENERAL INFORMATION:  
APPLICANT: Ni et al.  
TITLE OF INVENTION: 207 Human Secreted Proteins  
FILE REFERENCE: P2007P2  
CURRENT APPLICATION NUMBER: US/09/933,767  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: PCT/US01/05614  
PRIOR FILING DATE: 2001-02-21  
PRIOR APPLICATION NUMBER: 60/184,836  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: 60/193,170  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 09/205,258  
PRIOR FILING DATE: 1998-12-04  
PRIOR APPLICATION NUMBER: PCT/US98/11422  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/048,885  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/049,375  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,881  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,880  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,896  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/049,020  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,876

PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,895  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,884  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,894  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,971  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,964  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,882  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,899  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,893  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,900  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,901  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,892  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,915  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/049,019  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,970  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,972  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,916  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/049,373  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,875  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/049,374  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,917  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,949  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,974  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,883  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,897  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,898  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,962  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,963  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,877  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/048,878  
 PRIOR FILING DATE: 1997-06-06  
 PRIOR APPLICATION NUMBER: 60/068,054  
 PRIOR FILING DATE: 1997-12-18  
 PRIOR APPLICATION NUMBER: 60/068,064  
 PRIOR FILING DATE: 1997-12-18  
 PRIOR APPLICATION NUMBER: 60/068,053  
 PRIOR FILING DATE: 1997-12-18  
 PRIOR APPLICATION NUMBER: 60/070,923  
 PRIOR FILING DATE: 1997-12-18  
 PRIOR APPLICATION NUMBER: 60/073,160  
 PRIOR FILING DATE: 1998-01-30  
 PRIOR APPLICATION NUMBER: 60/073,159  
 PRIOR FILING DATE: 1998-01-30  
 PRIOR APPLICATION NUMBER: 60/073,165  
 PRIOR FILING DATE: 1998-01-30

PRIOR APPLICATION NUMBER: 60/073,164  
 PRIOR FILING DATE: 1998-01-30  
 PRIOR APPLICATION NUMBER: 60/085,925  
 PRIOR FILING DATE: 1998-05-18  
 PRIOR APPLICATION NUMBER: 60/085,921  
 PRIOR FILING DATE: 1998-05-18  
 PRIOR APPLICATION NUMBER: 60/085,923  
 PRIOR FILING DATE: 1998-05-18  
 PRIOR APPLICATION NUMBER: 60/085,922  
 PRIOR FILING DATE: 1998-05-18  
 PRIOR APPLICATION NUMBER: 60/092,921  
 PRIOR FILING DATE: 1998-07-15  
 PRIOR APPLICATION NUMBER: 60/094,657  
 PRIOR FILING DATE: 1998-07-30  
 NUMBER OF SEQ ID NOS: 1245  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 450  
 LENGTH: 101  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: SITE  
 LOCATION: (44)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (46)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (77)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (78)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (101)  
 OTHER INFORMATION: Xaa equals stop translation  
 US-09-933-767-450

Query Match 4.9%; Score 7; DB 10; Length 101;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 FLKHRI 19  
 DB 22 FLKHRI 28

RESULT 14  
 US-10-004-860-450  
 Sequence 450, Application US/10004860  
 Publication No. US20030065160A1  
 GENERAL INFORMATION:  
 APPLICANT: Young et al.  
 TITLE OF INVENTION: 207 Human Secreted Proteins  
 FILE REFERENCE: P2007P1  
 CURRENT APPLICATION NUMBER: US/10/004,860  
 CURRENT FILING DATE: 2001-12-07  
 Prior Application removed - See File Wrapper or Palm  
 NUMBER OF SEQ ID NOS: 1227  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 450  
 LENGTH: 101  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: SITE  
 LOCATION: (44)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (46)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 FEATURE:  
 NAME/KEY: SITE  
 LOCATION: (46)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids



NAME/KEY: SITE  
LOCATION: (77)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (78)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (101)  
OTHER INFORMATION: Xaa equals stop translation  
US-10-004-860-450

Query Match 4.9%; Score 7; DB 12; Length 101;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 FLKDHRI 19  
Db 22 FLKDHRI 28

RESULT 15  
US-10-023-282-450  
Sequence 450, Application US/10023282  
Publication No. US20030092893A1  
GENERAL INFORMATION:  
APPLICANT: Young et al.  
TITLE OF INVENTION: 207 Human Secreted Proteins  
FILE REFERENCE: P2007P1  
CURRENT APPLICATION NUMBER: US/10/023,282  
EARLIER FILING DATE: 2001-12-20  
EARLIER APPLICATION NUMBER: 1998-12-04  
EARLIER FILING DATE: 1998-06-04  
EARLIER APPLICATION NUMBER: PCT/US98/11422  
EARLIER FILING DATE: 1998-06-04  
EARLIER APPLICATION NUMBER: 60/048,885  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,375  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,881  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,880  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,896  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,020  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,876  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,893  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,900  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915

EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,916  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,373  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 450  
LENGTH: 101  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (44)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (77)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (78)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (101)  
OTHER INFORMATION: Xaa equals stop translation  
US-10-023-282-450

Query Match 4.9%; Score 7; DB 14; Length 101;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 FLKDRI 19  
|||||  
Db 22 FLKDRI 28

## RESULT 16

US-10-424-599-209010  
; Sequence 209010, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 209010  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_30764C.1.pep  
US-10-424-599-209010

Query Match 4.9%; Score 7; DB 12; Length 126;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 KIAKRN 118  
|||||  
Db 37 KIAKRN 43

RESULT 17  
US-10-437-963-116190  
; Sequence 116190, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 116190  
; LENGTH: 176  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_19714C.1.pep  
US-10-437-963-116190

Query Match 4.9%; Score 7; DB 16; Length 176;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPPA 9  
|||||  
Db 97 APTLPPA 103

RESULT 18  
US-10-282-122A-60941  
; Sequence 60941, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 60941  
; LENGTH: 179  
; TYPE: PRT  
; ORGANISM: Listeria monocytogenes  
US-10-282-122A-60941

Query Match 4.9%; Score 7; DB 12; Length 179;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 TAKVRR 133  
|||||  
Db 36 TAKVRR 42

RESULT 19  
US-10-282-122A-57125  
; Sequence 57125, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant

```

; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57125
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-10-282-122A-57125

Query Match
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 TAKKVR 133
DB 39 TAKKVR 45

RESULT 20
US-10-282-122A-57761
; Sequence 57761, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335

```

```

; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57761
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Enterococcus faecium
; US-10-282-122A-57761

Query Match
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 TAKKVR 133
DB 39 TAKKVR 45

RESULT 21
US-10-424-599-245457
; Sequence 245457, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 265684
; SEQ ID NO 245457
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_6367C.1.pep
; US-10-424-599-245457

Query Match
Best Local Similarity 100.0%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 120 KKKFEE 126
DB 43 KKKFEE 49

RESULT 22
US-09-815-242-10698
; Sequence 10698, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.

```

APPLICANT: Carr, Grant J.  
 APPLICANT: Yamamoto, Robert T.  
 TITLE OF INVENTION: Identification of Essential Genes in  
 FILE REFERENCE: ELITRA.011A  
 CURRENT APPLICATION NUMBER: US/09/815,242  
 CURRENT FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 NUMBER OF SEQ ID NOS: 14110  
 SOFTWARE: FASTSEQ for Windows Version 4.0  
 SEQ ID NO: 10698  
 LENGTH: 194  
 TYPE: PR1  
 ORGANISM: Enterococcus faecalis  
 US-09-815-242-10698

Query Match 4.9%; Score 7; DB 9; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 TAKVVR 133  
 DB 51 TAKVVR 57

RESULT 23  
 US-10-335-977-7963  
 Sequence 7963, Application US/10335977  
 Publication No. US20040052799A1  
 GENERAL INFORMATION:  
 APPLICANT: DOUGLAS SMITH et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
 RELATING TO HELICOBACTER PYLORI FOR  
 DIAGNOSTICS AND THERAPEUTICS  
 NUMBER OF SEQUENCES: 10031  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHYE & COCKFIELD  
 STREET: 28 State Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109-1875  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: CD-ROM ISO9660  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: Windows NT 4.0  
 SOFTWARE: UNIX  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/335,977  
 FILING DATE: 30-Dec-2002  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/993,002  
 FILING DATE: 17-Dec-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Mandragouras, Amy E.  
 REGISTRATION NUMBER: 36,207  
 REFERENCE/DOCKET NUMBER: GTN-018  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214  
 INFORMATION FOR SEQ ID NO: 7963:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 199 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: YES  
 ORIGINAL SOURCE:  
 ORGANISM: Helicobacter pylori  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (8) LOCATION 1...199  
 SEQUENCE DESCRIPTION: SEQ ID NO: 7963:  
 US-10-335-977-7963

Query Match 4.9%; Score 7; DB 12; Length 199;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 ELTLGRF 101  
 DB 98 ELTLGRF 104

RESULT 24  
 US-10-335-977-7964  
 Sequence 7964, Application US/10335977  
 Publication No. US20040052799A1  
 GENERAL INFORMATION:  
 APPLICANT: DOUGLAS SMITH et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
 RELATING TO HELICOBACTER PYLORI FOR  
 DIAGNOSTICS AND THERAPEUTICS  
 NUMBER OF SEQUENCES: 10031  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHYE & COCKFIELD  
 STREET: 28 State Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109-1875  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: CD-ROM ISO9660  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: Windows NT 4.0  
 SOFTWARE: UNIX  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/335,977  
 FILING DATE: 30-Dec-2002  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/993,002  
 FILING DATE: 17-Dec-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Mandragouras, Amy E.  
 REGISTRATION NUMBER: 36,207  
 REFERENCE/DOCKET NUMBER: GTN-018  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)227-7400  
 TELEFAX: (617)742-4214  
 INFORMATION FOR SEQ ID NO: 7964:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 203 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: YES  
 ORIGINAL SOURCE:  
 ORGANISM: Helicobacter pylori  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (8) LOCATION 1...203  
 SEQUENCE DESCRIPTION: SEQ ID NO: 7964:  
 US-10-335-977-7964

US-10-335-977-7964

Query Match  
Best Local Similarity 100.0%; Score 7; DB 12; Length 203;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 ELTIGER 101  
DB 102 ELTIGER 108

RESULT 25

US-10-424-599-243648  
; Sequence 243648, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 243648  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_62040C.1.pep  
US-10-424-599-243648

Query Match  
Best Local Similarity 100.0%; Score 7; DB 12; Length 259;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 RAKXKIA 114  
DB 70 RAKXKIA 76

RESULT 26

US-10-424-599-147883  
; Sequence 147883, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 147883  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(311)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_104560C.1.pep  
US-10-424-599-147883

Query Match  
Best Local Similarity 100.0%; Score 7; DB 12; Length 311;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KYRAIE 136

DB 137 KYRAIE 143

RESULT 27

US-10-425-114-50128  
; Sequence 50128, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 50128  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700682056\_FLI.pep  
US-10-425-114-50128

Query Match  
Best Local Similarity 100.0%; Score 7; DB 12; Length 311;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KYRAIE 136  
DB 138 KYRAIE 144

RESULT 28

US-10-369-493-13992  
; Sequence 13992, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xiandeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 13992  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Pseudomonas fluorescens  
US-10-369-493-13992

Query Match  
Best Local Similarity 100.0%; Score 7; DB 15; Length 382;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 AIEQLAA 140  
DB 18 AIEQLAA 24

```

RESULT 29
US-10-424-599-154013
; Sequence 154013, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154013
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110095C.1.pep
US-10-424-599-154013

```

```

Query Match          4.9%; Score 7; DB 12; Length 388;
Best Local Similarity 100.0%; Pred.No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      124 FEETAKK 130
DB      187 FEETAKK 193

```

```

RESULT 30
US-10-424-599-248527
; Sequence 248527, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 248527
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(415)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_6644C.1.pep
US-10-424-599-248527

```

```

Query Match          4.9%; Score 7; DB 12; Length 415;
Best Local Similarity 100.0%; Pred.No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      128 AKVRRRA 134
DB      159 AKVRRRA 165

```

```

RESULT 31
US-10-421-654-28
; Sequence 28, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
US-10-421-654-28

```

```

Query Match          4.9%; Score 7; DB 15; Length 428;
Best Local Similarity 100.0%; Pred.No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      36 ERMAENG 42
DB      373 ERMAENG 379

```

```

RESULT 32
US-10-421-654-44
; Sequence 44, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
US-10-421-654-44

```

```

Query Match          4.9%; Score 7; DB 15; Length 428;
Best Local Similarity 100.0%; Pred.No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      36 ERMAENG 42
DB      373 ERMAENG 379

```

```

RESULT 33
US-10-369-493-10000
; Sequence 10000, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.

```

APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
FILE REFERENCE: 38-10(52052)B  
CURRENT FILING DATE: 2003-02-28  
CURRENT APPLICATION NUMBER: US/10/369,493  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 10000  
LENGTH: 475  
TYPE: PRT  
ORGANISM: magnetite-containing magnetic coccus  
US-10-369-493-10000

Query Match 4.9%; Score 7; DB 15; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 VRRALIQ 137  
DB 284 VRRALIQ 290

## RESULT 34

US-10-437-963-192686  
Sequence 192686, Application US/10437963  
Publication NO. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 192686  
LENGTH: 610  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(610)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_88895C.1.pep  
US-10-437-963-192686

Query Match 4.9%; Score 7; DB 16; Length 610;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPPA 9  
DB 81 APTLPPA 87

## RESULT 35

US-10-369-493-21957  
Sequence 21957, Application US/10369493  
Publication NO. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
FILE REFERENCE: 38-10(52052)B  
CURRENT FILING DATE: 2003-02-28  
CURRENT APPLICATION NUMBER: US/10/369,493  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 21957  
LENGTH: 744  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-10-369-493-21957

Query Match 4.9%; Score 7; DB 15; Length 744;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 EELTLGE 100  
DB 587 EELTLGE 593

## RESULT 36

US-10-437-963-201149  
Sequence 201149, Application US/10437963  
Publication NO. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 201149  
LENGTH: 751  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(751)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_96551C.1.pep  
US-10-437-963-201149

Query Match 4.9%; Score 7; DB 16; Length 751;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 HKHSSG 83  
DB 175 HKHSSG 181

## RESULT 37

US-10-437-963-180519  
Sequence 180519, Application US/10437963  
Publication NO. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua

```
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 180519
LENGTH: 802
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_77882C.1.pep
US-10-437-963-180519
```

```
Query Match      4.9%; Score 7; DB 16; Length 802;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      134 ATEQLAA 140
DB      301 ATEQLAA 307
```

```
RESULT 38
US-10-128-714-3257
Sequence 3257, Application US/10128714
Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Wengqi
APPLICANT: Tishoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroshkin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3257
LENGTH: 838
TYPE: PRT
ORGANISM: Aspergillus fumigatus
US-10-128-714-3257
```

```
Query Match      4.9%; Score 7; DB 14; Length 838;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      94 EELTIGE 100
DB      11 EELTIGE 17
```

```
RESULT 39
US-09-932-183A-2
```

```
Sequence 2, Application US/0932183A
Patent No. US20020127641A1
GENERAL INFORMATION:
APPLICANT: Estell, David A.
TITLE OF INVENTION: Proteases From Gram-Positive Organisms
FILE REFERENCE: GC394C1-US
CURRENT APPLICATION NUMBER: US/09/932,183A
CURRENT FILING DATE: 2001-08-17
PRIOR APPLICATION NUMBER: US 09/308,375
PRIOR FILING DATE: 1998-05-14
PRIOR APPLICATION NUMBER: PCT/US98/18828
PRIOR FILING DATE: 1998-09-08
PRIOR APPLICATION NUMBER: EP9719636.4
PRIOR FILING DATE: 1997-09-15
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 2285
TYPE: PRT
ORGANISM: Bacillus subtilis
US-09-932-183A-2
```

```
Query Match      4.9%; Score 7; DB 9; Length 2285;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      132 RRAIEQL 138
DB      184 RRAIEQL 190
```

```
RESULT 40
US-10-083-768-130
Sequence 130, Application US/10083768
Publication No. US20030158116A1
GENERAL INFORMATION:
APPLICANT: Dower, William J.
APPLICANT: Barrett, Ronald W.
APPLICANT: Cwiria, Steven E.
APPLICANT: Duffin, David J.
APPLICANT: Gates, Christian
APPLICANT: Haselden, Sherill S.
APPLICANT: Matheakis, Larry C.
APPLICANT: Schatz, Peter J.
APPLICANT: Wagstrom, Christopher R.
APPLICANT: Wrighton, Nicholas C.
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
FILE REFERENCE: THROMBOPOIETIN RECEPTOR
NUMBER OF SEQUENCES: 232
CORRESPONDENCE ADDRESS:
ADDRESS: Glaxo Wellcome
STREET: Five Moore Drive, P.O. Box 13398
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/083,768
FILING DATE: 27-Feb-2002
ATTORNEY/AGENT INFORMATION:
NAME: Hrubiec, Robert T.
REGISTRATION NUMBER: 36,392
REFERENCE/DOCKET NUMBER: PK3065USM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-248-1000
INFORMATION FOR SEQ ID NO: 130:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
```



TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 130:  
US-10-083-768-130

Query Match 4.2%; Score 6; DB 14; Length 12;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 TLGEFL 102  
Db 2 TLGEFL 7

RESULT 41  
US-09-939-980-380  
Sequence 380, Application US/09939980  
Patent No. US2002008234A1  
GENERAL INFORMATION:  
APPLICANT: Black, Michael  
Burnham, Martin  
Hodgson, John  
Knowles, David  
Loretto, Michael  
Nicholas, Richard  
Prett, Julie  
Reichard, Richard  
Rosenberg, Martin  
Ward, Judith  
TITLE OF INVENTION: No. US2002008234A1el Prokaryotic Polynucleotides,  
Polypeptides and Their Uses  
NUMBER OF SEQUENCES: 534  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,980  
FILING DATE: 27-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/936,165  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimm, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50549  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 380:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 380:  
US-09-939-980-380

Query Match 4.2%; Score 6; DB 9; Length 30;  
Best Local Similarity 100.0%; Pred. No. 2,1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 109 AKNKIA 114  
Db 9 AKNKIA 14

RESULT 42  
US-09-989-919-109  
Sequence 109, Application US/09989919  
Patent No. US2002016344A1  
GENERAL INFORMATION:  
APPLICANT: Macina, Roberto  
APPLICANT: Recipon, Heeye  
APPLICANT: Pluta, Jason  
APPLICANT: Ghosh, Malavika  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Prot  
FILE REFERENCE: DEX-0289  
CURRENT APPLICATION NUMBER: US/09/989,919  
CURRENT FILING DATE: 2001-11-21  
PRIOR APPLICATION NUMBER: 60/252,505  
PRIOR FILING DATE: 2000-11-22  
NUMBER OF SEQ ID NOS: 124  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 109  
LENGTH: 36  
TYPE: PRT  
ORGANISM: Homo sapien  
US-09-989-919-109

Query Match 4.2%; Score 6; DB 9; Length 36;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LSVKKQ 92  
Db 21 LSVKKQ 26

RESULT 43  
US-10-424-599-262708  
Sequence 262708, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J  
APPLICANT: Kovalic, David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
Fate Reference: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 262708  
LENGTH: 42  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_79247C.1.pep  
US-10-424-599-262708

Query Match 4.2%; Score 6; DB 12; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 TPKXWP 26  
Db 36 TPKXWP 41

Query Match 4.2%; Score 6; DB 9; Length 30;  
Best Local Similarity 100.0%; Pred. No. 2,1e+02;

RESULT 44  
US-10-424-599-276718  
; Sequence 276718, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 276718  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_91897C.1.pep  
US-10-424-599-276718

Query Match  
Best Local Similarity 4.2%; Score 6; DB 12; Length 52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 TNNKKX 122  
DB 11 TNNKKX 16

RESULT 45  
US-10-425-114-68762  
; Sequence 68762, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 68762  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17127B08\_FLI.pep  
US-10-425-114-68762

Query Match  
Best Local Similarity 4.2%; Score 6; DB 12; Length 52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 NNRKKE 123  
DB 16 NNRKKE 21

RESULT 46  
US-09-860-352A-11  
; Sequence 11, Application US/09860352A  
; Patent No. US20020132785A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Curtis, Rory

APPLICANT: Welch, Nadine  
; TITLE OF INVENTION: 13305 NOVEL PROTEIN KINASE MOLECULES AND  
; TITLE OF INVENTION: USES THEREFOR  
; FILE REFERENCE: 38155-20016.00  
; CURRENT APPLICATION NUMBER: US/09/860,352A  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: US 60/205,301  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 55  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Consensus amino acid  
US-09-860-352A-11

Query Match  
Best Local Similarity 4.2%; Score 6; DB 9; Length 55;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LPPAWQ 11  
DB 41 LPPAWQ 46

RESULT 47  
US-10-424-599-152902  
; Sequence 152902, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 152902  
; LENGTH: 55  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_109094C.1.pep  
US-10-424-599-152902

Query Match  
Best Local Similarity 4.2%; Score 6; DB 12; Length 55;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPP 8  
DB 38 APTLPP 43

RESULT 48  
US-10-410-764-98  
; Sequence 98, Application US/10410764  
; Publication No. US2004000564A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Wachselt, Kyle J.  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Welch, Nadine S.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Tsai, Peng-Ying  
; APPLICANT: Kapeller-Libermann, Rosana

APPLICANT: Carroll, Joseph M.  
TITLE OF INVENTION: 26199, 33530, 33949, 47148, 50226,  
TITLE OF INVENTION: 58764, 62113, 32144, 32235, 23565, 13305, 14911, 86216,  
TITLE OF INVENTION: 25206 AND 8843 MOLECULES AND USES THEREFOR  
FILE REFERENCE: MPIO3-0520MIM  
CURRENT APPLICATION NUMBER: US/10/410,764  
CURRENT FILING DATE: 2003-04-10  
PRIOR APPLICATION NUMBER: US 09/924,358  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: US 60/229,300  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: US 10/350,553  
PRIOR FILING DATE: 2003-01-24  
PRIOR APPLICATION NUMBER: US 60/351,572  
PRIOR FILING DATE: 2002-01-24  
PRIOR APPLICATION NUMBER: US 09/966,614  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: US 60/238,054  
PRIOR FILING DATE: 2000-10-05  
PRIOR APPLICATION NUMBER: US 10/281,094  
PRIOR FILING DATE: 2002-10-25  
PRIOR APPLICATION NUMBER: US 60/347,815  
PRIOR FILING DATE: 2001-10-29  
PRIOR APPLICATION NUMBER: US 10/076,535  
PRIOR FILING DATE: 2002-02-15  
PRIOR APPLICATION NUMBER: US 60/269,440  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 136  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 98  
LENGTH: 55  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Amino acid consensus sequence  
US-10-410-764-98

Query Match 4.2%; Score 6; DB 15; Length 55;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6 LPPAMQ 11  
DB 41 LPPAMQ 46  
RESULT 49  
US-10-424-599-158384  
Sequence 158384, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 158384  
LENGTH: 56  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MFT3847\_114039C.1.pdp  
US-10-424-599-158384  
Query Match 4.2%; Score 6; DB 12; Length 56;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPP 8  
DB 44 APTLPP 49  
RESULT 50  
US-10-424-599-252199  
Sequence 252199, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 252199  
LENGTH: 57  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MFT3847\_69764C.1.pdp  
US-10-424-599-252199  
Query Match 4.2%; Score 6; DB 12; Length 57;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 110 KNKIAK 115  
DB 11 KNKIAK 16

RESULT 51  
US-10-125-258-62  
Sequence 62, Application US/10125258  
Publication No. US20030028920A1  
GENERAL INFORMATION:  
APPLICANT: Altier, Daniel J.  
APPLICANT: Hermann, Rafael  
APPLICANT: Lu, Albert L.  
APPLICANT: McCutchen, Billy F.  
APPLICANT: Prenatal, James K.  
APPLICANT: Weaver, Janine L.  
APPLICANT: Wong, James F. H.  
TITLE OF INVENTION: Antimicrobial Polypeptides and Their  
FILE REFERENCE: 35718/246215  
CURRENT APPLICATION NUMBER: US/10/125,258  
CURRENT FILING DATE: 2002-04-18  
PRIOR APPLICATION NUMBER: 60/285,355  
PRIOR FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 127  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 62  
LENGTH: 57  
TYPE: PRT  
ORGANISM: Helicoverpa zea  
US-10-125-258-62

Query Match 4.2%; Score 6; DB 14; Length 57;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 61 FKELEG 66  
DB 25 FKELEG 30

RESULT 52  
US-10-424-599-167067  
; Sequence 167067, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 167067  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_121876C.1.pep  
US-10-424-599-167067

Query Match 4.2%; Score 6; DB 12; Length 58;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 KKKKEE 125  
DB 26 KKKKEE 31

RESULT 53  
US-10-424-599-271511  
; Sequence 271511, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 271511  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_87192C.1.pep  
US-10-424-599-271511

Query Match 4.2%; Score 6; DB 12; Length 58;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 FFCFKE 63  
DB 43 FFCFKE 48

RESULT 54  
US-10-437-963-183020  
; Sequence 183020, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua

APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 183020  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_8014C.1.pep  
US-10-437-963-183020

Query Match 4.2%; Score 6; DB 16; Length 58;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FLSVKK 91  
DB 13 FLSVKK 18

RESULT 55  
US-10-424-599-194266  
; Sequence 194266, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 194266  
; LENGTH: 62  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_1744C.1.pep  
US-10-424-599-194266

Query Match 4.2%; Score 6; DB 12; Length 62;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PFLKDH 17  
DB 52 PFLKDH 57

RESULT 56  
US-10-424-599-237412  
; Sequence 237412, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599

```

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 237412
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(63)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_56408C.1.pcp
US-10-424-599-237412

Query Match
Best Local Similarity 100.0%; Score 6; DB 12; Length 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 NNNKKE 123
Db 12 NNNKKE 17

RESULT 57
US-10-125-258-61
; Sequence 61, Application US/10125258
; Publication No. US20030028920A1
; GENERAL INFORMATION:
; APPLICANT: Altier, Daniel J.
; APPLICANT: Herrmann, Rafael
; APPLICANT: Lu, Albert L.
; APPLICANT: McCutchen, Billy F.
; APPLICANT: Presnell, James K.
; APPLICANT: Weaver, Janine L.
; APPLICANT: Wong, James F. H.
; TITLE OF INVENTION: Antimicrobial Polypeptides and Their
; TITLE OF INVENTION: Uses
; FILE REFERENCE: 35718/246215
; CURRENT APPLICATION NUMBER: US/10/125,258
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 60/285,355
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Helicoverpa zea
US-10-125-258-61

Query Match
Best Local Similarity 100.0%; Score 6; DB 14; Length 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 FKELEG 66
Db 25 FKELEG 30

RESULT 58
US-10-425-114-54599
; Sequence 54599, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53313)B
```

```

; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54599
; LENGTH: 64
; TYPE: PRT
; ORGANISM: zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMOTECOSINTE105D05_FLI.pcp
US-10-425-114-54599

Query Match
Best Local Similarity 100.0%; Score 6; DB 12; Length 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 VRRATE 136
Db 26 VRRATE 31

RESULT 59
US-09-764-881-122
; Sequence 122, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 122
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-881-122

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KKEFEE 126
Db 20 KKEFEE 25

RESULT 60
US-09-764-875-1113
; Sequence 1113, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1113
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-875-1113

Query Match
Best Local Similarity 100.0%; Score 6; DB 11; Length 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KKEFEE 126
Db 20 KKEFEE 25
```

Db 20 KKEEFE 25

RESULT 61

US-10-424-599-281627  
; Sequence 281627, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 281627  
; LENGTH: 65  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(65)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_9632C.1.pep  
US-10-424-599-281627

Query Match

Best Local Similarity 4.2%; Score 6; DB 12; Length 65;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FLSVKK 91

Db 32 FLSVKK 37

RESULT 62

US-09-764-881-122  
; Sequence 122, Application US/09764881  
; Publication No. US20020086821A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT207  
; CURRENT APPLICATION NUMBER: US/09/764,881  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 122  
; LENGTH: 65  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-881-122

Query Match

Best Local Similarity 4.2%; Score 6; DB 12; Length 65;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KKEEFE 126

Db 20 KKEEFE 25

RESULT 63

US-10-242-747-122  
; Sequence 122, Application US/10242747  
; Publication No. US20040005577A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PT207C1  
; CURRENT APPLICATION NUMBER: US/10/242,747  
; CURRENT FILING DATE: 2002-09-13  
; Prior APPLICATION NUMBER: 09/764,881  
; Prior FILING DATE: 2001-01-17  
; Prior APPLICATION NUMBER: 60/179,065  
; Prior FILING DATE: 2000-01-31  
; Prior APPLICATION NUMBER: 60/180,628  
; Prior FILING DATE: 2000-02-04  
; Prior APPLICATION NUMBER: 60/214,886  
; Prior FILING DATE: 2000-06-28  
; Prior APPLICATION NUMBER: 60/217,487  
; Prior FILING DATE: 2000-07-11  
; Prior APPLICATION NUMBER: 60/225,758  
; Prior FILING DATE: 2000-08-14  
; Prior APPLICATION NUMBER: 60/220,963  
; Prior FILING DATE: 2000-07-26  
; Prior APPLICATION NUMBER: 60/217,496  
; Prior FILING DATE: 2000-07-11  
; Prior APPLICATION NUMBER: 60/225,447  
; Prior FILING DATE: 2000-08-14  
; Prior APPLICATION NUMBER: 60/218,290  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 122  
; LENGTH: 65  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-242-747-122

Query Match

Best Local Similarity 4.2%; Score 6; DB 15; Length 65;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KKEEFE 126

Db 20 KKEEFE 25

RESULT 64

US-10-158-057-235  
; Sequence 235, Application US/10158057  
; Publication No. US20040014039A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P205C1  
; CURRENT APPLICATION NUMBER: US/10/158,057  
; CURRENT FILING DATE: 2002-06-12  
; Prior application data removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 384  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 235  
; LENGTH: 65  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-158-057-235

Query Match

Best Local Similarity 4.2%; Score 6; DB 15; Length 65;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KKEEFE 126

Db 20 KKEEFE 25

RESULT 65

US-10-424-599-199019  
; Sequence 199019, Application US/10424599

```

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 199019
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_21739C.1.pep
US-10-424-599-199019

```

```

Query Match      4.2%; Score 6; DB 12; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      84 CAPLVS 89
Db      60 CAPLVS 65

```

```

RESULT 66
US-10-424-599-167398
; Sequence 167398, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 167398
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_122178C.1.pep
US-10-424-599-167398

```

```

Query Match      4.2%; Score 6; DB 12; Length 67;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      15 KDHRS 20
Db      50 KDHRS 55

```

```

RESULT 67
US-10-424-599-153000
; Sequence 153000, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B

```

```

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 153000
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_109183C.1.pep
US-10-424-599-153000

```

```

Query Match      4.2%; Score 6; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      86 FLVSVK 91
Db      38 FLVSVK 43

```

```

RESULT 68
US-10-437-963-141151
; Sequence 141151, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 141151
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(74)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_42283C.1.pep
US-10-437-963-141151

```

```

Query Match      4.2%; Score 6; DB 16; Length 74;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      7 PPAWCP 12
Db      62 PPAWCP 67

```

```

RESULT 69
US-10-282-122A-45175
; Sequence 45175, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John

```

```

; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45175
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-45175

Query Match          4.2%; Score 6; DB 12; Length 75;
Best Local Similarity 100.0%; Pred.No.4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      125 EETAKK 130
Db      14 EETAKK 19

RESULT 70
US-10-437-963-182219
; Sequence 182219, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 182219
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(76)
; OTHER INFORMATION: unsure at all Xaa locations
```

```

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_79427C.1.pep
US-10-437-963-182219

Query Match          4.2%; Score 6; DB 16; Length 76;
Best Local Similarity 100.0%; Pred.No.4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 PTLPPA 9
Db      47 PTLPPA 52

RESULT 71
US-10-424-599-144105
; Sequence 144105, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 144105
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(77)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101139C.1.pep
US-10-424-599-144105

Query Match          4.2%; Score 6; DB 12; Length 77;
Best Local Similarity 100.0%; Pred.No.4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 APTLPP 8
Db      65 APTLPP 70

RESULT 72
US-10-424-599-199222
; Sequence 199222, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 199222
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(82)
; OTHER INFORMATION: Clone ID: PAT_MRT3847_21921C.1.pep
US-10-424-599-199222

Query Match          4.2%; Score 6; DB 12; Length 82;
```



Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 LDRERA 109  
|||||  
Db 47 LDRERA 52

## RESULT 73

US-10-424-599-152919  
; Sequence 152919, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 152919  
; LENGTH: 84  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_109109C.1.pep  
US-10-424-599-152919

Query Match  
Best Local Similarity 100.0%; Score 6; DB 12; Length 84;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 TPERMA 39  
|||||  
Db 73 TPERMA 78

## RESULT 74

US-10-424-599-232417  
; Sequence 232417, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 232417  
; LENGTH: 84  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_51897C.1.pep  
US-10-424-599-232417

Query Match  
Best Local Similarity 100.0%; Score 6; DB 12; Length 84;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPP 8  
|||||  
Db 52 APTLPP 57

## RESULT 75

US-10-424-599-216095  
; Sequence 216095, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 216095  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_37163C.1.pep  
US-10-424-599-216095

Query Match  
Best Local Similarity 100.0%; Score 6; DB 12; Length 85;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RERAXN 111  
|||||  
Db 11 RERAXN 16

Search completed: August 11, 2004, 14:30:41  
Job time : 48 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 14:22:47 ; Search time 19 Seconds

(without alignments)  
385.836 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 142

Sequence: 1 MGAPLPPANQPIKDHRI.....EPEETAKKVRALQLAAMD 142

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625871 residues

Word size: 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCUTS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfillset1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142	100.0	142	US-08-975-080-34	Sequence 34, Appl1
2	142	100.0	142	US-09-283-144-3	Sequence 3, Appl1
3	37	26.1	37	US-09-283-144-15	Sequence 15, Appl1
4	20	14.1	20	US-08-975-080-4	Sequence 4, Appl1
5	17	12.0	17	US-08-975-080-3	Sequence 3, Appl1
6	17	12.0	17	US-09-515-514-1	Sequence 1, Appl1
7	9	6.3	9	US-09-515-514-12	Sequence 12, Appl1
8	9	6.3	9	US-09-515-514-13	Sequence 13, Appl1
9	9	6.3	9	US-09-205-258-450	Sequence 450, App
10	7	4.9	130	US-09-732-210-281	Sequence 281, App
11	7	4.9	167	US-09-107-532A-5332	Sequence 5332, Ap
12	7	4.9	186	US-09-107-532A-5027	Sequence 5027, Ap
13	7	4.9	208	US-09-134-000C-4317	Sequence 4317, Ap
14	7	4.9	258	US-09-134-000C-3780	Sequence 3780, Ap
15	7	4.9	307	US-09-475-316A-70	Sequence 70, Appl1
16	7	4.9	307	US-09-704-640-70	Sequence 70, Appl1
17	7	4.9	309	US-09-475-316A-72	Sequence 72, Appl1
18	7	4.9	309	US-09-704-640-72	Sequence 72, Appl1
19	7	4.9	313	US-09-475-316A-62	Sequence 62, Appl1
20	7	4.9	313	US-09-704-640-62	Sequence 62, Appl1
21	7	4.9	314	US-09-475-316A-66	Sequence 66, Appl1
22	7	4.9	314	US-09-475-316A-118	Sequence 118, App
23	7	4.9	314	US-09-704-640-66	Sequence 66, Appl1
24	7	4.9	314	US-09-704-640-118	Sequence 118, App
25	7	4.9	418	US-09-252-991A-25075	Sequence 25075, A
26	7	4.9	580	US-09-107-532A-5734	Sequence 5734, Ap
27	7	4.9	2285	US-09-308-375-2	Sequence 2, Appl1

28	6	4.2	11	3	US-08-928-383B-5	Sequence 5, Appl1
29	6	4.2	12	2	US-08-764-640-130	Sequence 130, App
30	6	4.2	12	3	US-08-973-225-130	Sequence 130, App
31	6	4.2	12	3	US-09-244-298A-130	Sequence 130, App
32	6	4.2	12	3	US-09-516-704-130	Sequence 130, App
33	6	4.2	12	4	US-09-549-090-130	Sequence 130, App
34	6	4.2	12	4	US-09-832-230A-130	Sequence 130, App
35	6	4.2	24	1	US-08-185-433-13	Sequence 13, App
36	6	4.2	30	4	US-08-936-165A-380	Sequence 380, App
37	6	4.2	58	4	US-09-621-976-6592	Sequence 6992, App
38	6	4.2	80	4	US-09-328-352-4702	Sequence 4702, App
39	6	4.2	93	4	US-09-107-532A-3689	Sequence 3689, App
40	6	4.2	95	4	US-09-252-991A-19558	Sequence 19558, A
41	6	4.2	97	4	US-09-134-001C-5653	Sequence 5653, App
42	6	4.2	99	4	US-09-621-976-7607	Sequence 7607, App
43	6	4.2	130	4	US-09-489-039A-12130	Sequence 12130, A
44	6	4.2	133	4	US-09-252-991A-17914	Sequence 17914, A
45	6	4.2	134	4	US-09-800-729-141	Sequence 141, App
46	6	4.2	135	4	US-09-328-352-4961	Sequence 4961, App
47	6	4.2	141	4	US-09-732-210-545	Sequence 545, App
48	6	4.2	148	1	US-07-998-003A-36	Sequence 36, Appl1
49	6	4.2	148	1	US-08-453-274B-36	Sequence 36, Appl1
50	6	4.2	148	1	US-08-453-695A-36	Sequence 36, Appl1
51	6	4.2	148	1	US-08-468-347-22	Sequence 22, Appl1
52	6	4.2	148	1	US-08-268-161A-36	Sequence 36, Appl1
53	6	4.2	148	2	US-08-467-389-22	Sequence 22, Appl1
54	6	4.2	148	2	US-08-779-379-22	Sequence 22, Appl1
55	6	4.2	148	2	US-08-469-218-22	Sequence 22, Appl1
56	6	4.2	148	2	US-08-453-702A-36	Sequence 36, Appl1
57	6	4.2	148	3	US-09-228-152-22	Sequence 22, Appl1
58	6	4.2	148	3	US-09-099-693-36	Sequence 36, Appl1
59	6	4.2	148	5	PCT-US93-12588-36	Sequence 36, Appl1
60	6	4.2	148	5	PCT-US93-08071-36	Sequence 36, Appl1
61	6	4.2	153	1	US-08-226-264-26	Sequence 26, Appl1
62	6	4.2	153	4	US-09-738-946-4	Sequence 4, Appl1
63	6	4.2	155	4	US-09-732-210-1644	Sequence 1644, App
64	6	4.2	163	4	US-09-134-000C-4018	Sequence 4018, App
65	6	4.2	175	2	US-08-809-267-5	Sequence 5, Appl1
66	6	4.2	175	5	PCT-US95-13622A-5	Sequence 5, Appl1
67	6	4.2	181	4	US-09-252-991A-16898	Sequence 16898, A
68	6	4.2	182	4	US-09-328-352-4216	Sequence 4216, App
69	6	4.2	182	4	US-09-107-532A-6802	Sequence 6802, App
70	6	4.2	204	4	US-09-328-352-6654	Sequence 6654, App
71	6	4.2	207	4	US-09-328-352-6654	Sequence 6654, App
72	6	4.2	248	4	US-09-252-991A-32920	Sequence 32920, A
73	6	4.2	251	4	US-09-252-991A-32920	Sequence 32920, A
74	6	4.2	252	4	US-09-543-681A-6028	Sequence 6028, App
75	6	4.2	260	4	US-09-252-991A-17964	Sequence 17964, A
76	6	4.2	263	4	US-09-489-039A-10079	Sequence 10079, App
77	6	4.2	265	4	US-09-543-681A-8110	Sequence 8110, App
78	6	4.2	267	4	US-09-543-681A-8110	Sequence 8110, App
79	6	4.2	286	4	US-09-172-952-31	Sequence 31, Appl1
80	6	4.2	290	4	US-09-134-000C-5874	Sequence 5874, App
81	6	4.2	301	4	US-09-352-991A-28771	Sequence 28771, App
82	6	4.2	308	4	US-09-489-039A-11473	Sequence 11473, A
83	6	4.2	320	4	US-09-543-681A-5787	Sequence 5787, App
84	6	4.2	325	3	US-09-108-020-49	Sequence 49, Appl1
85	6	4.2	330	4	US-09-679-279-10	Sequence 10, Appl1
86	6	4.2	333	4	US-09-134-000C-5637	Sequence 5637, App
87	6	4.2	335	3	US-09-446-504-80	Sequence 80, Appl1
88	6	4.2	335	4	US-09-712-266-80	Sequence 80, Appl1
89	6	4.2	340	4	US-09-543-681A-4668	Sequence 4668, App
90	6	4.2	342	1	US-08-854-196-2	Sequence 2, Appl1
91	6	4.2	342	3	US-09-064-033-2	Sequence 2, Appl1
92	6	4.2	342	4	US-09-118-464-3	Sequence 3, Appl1
93	6	4.2	342	4	US-09-291-046-2	Sequence 2, Appl1
94	6	4.2	348	1	US-08-843-521-6	Sequence 6, Appl1
95	6	4.2	348	2	US-08-853-041-2	Sequence 2, Appl1
96	6	4.2	348	3	US-09-159-417-2	Sequence 2, Appl1
97	6	4.2	358	3	US-09-012-871-6	Sequence 6, Appl1
98	6	4.2	355	4	US-08-630-915A-192	Sequence 192, App
99	6	4.2	361	3	US-09-041-718-4	Sequence 4, Appl1
100	6	4.2	364	4	US-09-134-000C-4030	Sequence 4030, App

## ALIGNMENTS

RESULT 1

US-08-975-080-34

Sequence 34, Application US/08975080

Patent No. 6245523

GENERAL INFORMATION:

APPLICANT: Altieri, Dario C.

TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS

TITLE OF INVENTION: CELLULAR APOPTOSIS, AND ITS MODULATION

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESS: MORGAN, LEWIS &amp; BOCKIUS LLP

STREET: 1800 M Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036-5869

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/975,080

FILING DATE: 20-NOV-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/031,435

FILING DATE: 20-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Adler, Reid G.

REGISTRATION NUMBER: 30,988

REFERENCE/DOCKET NUMBER: 044574-5022-01-WO

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-467-7000

TELEFAX: 202-467-7176

INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:

LENGTH: 142 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-975-080-34

Query Match 100.0%; Score 142; DB 3; Length 142;

Best Local Similarity 100.0%; Pred. No. 1.1e-138; Indels 0; Gaps 0;

Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGAPLPPAMQFPLKXDRISTFKNMPFLGCACTPERMAAGFIHCPTENEPDLAQCFFC 60

Db 1 MGAPLPPAMQFPLKXDRISTFKNMPFLGCACTPERMAAGFIHCPTENEPDLAQCFFC 60

Qy 61 FKELEGEPPDDPIEHHKHSSGCAFLSVKQFEEITLGEFLKDRERAKNKIAKETNNK 120

Db 61 FKELEGEPPDDPIEHHKHSSGCAFLSVKQFEEITLGEFLKDRERAKNKIAKETNNK 120

Qy 121 KKEFEETAKKVRRAIEQLAAMD 142

Db 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 2

US-09-283-144-3

Sequence 3, Application US/09283144

Patent No. 6346389

GENERAL INFORMATION:

APPLICANT: Yale University

TITLE OF INVENTION: Method for Selectively Modulating the Interactions

TITLE OF INVENTION: between Survivin and Tubulin

FILE REFERENCE: 44574-5033-US  
CURRENT APPLICATION NUMBER: US/09/283,144  
CURRENT FILING DATE: 1999-04-01  
EARLIER APPLICATION NUMBER: US 60/080,288  
EARLIER FILING DATE: 1998-04-01  
EARLIER APPLICATION NUMBER: US 08/975,080  
EARLIER FILING DATE: 1997-11-20  
EARLIER APPLICATION NUMBER: PCT/US97/21880  
EARLIER FILING DATE: 1997-11-20  
EARLIER APPLICATION NUMBER: US 60/031,435  
EARLIER FILING DATE: 1996-11-20  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Amino acid sequence of Survivin  
US-09-283-144-3

Query Match 100.0%; Score 142; DB 4; Length 142;

Best Local Similarity 100.0%; Pred. No. 1.1e-136; Indels 0; Gaps 0;

Matches 142; Conservative 0; Mismatches 0; Indels 0;

Qy 1 MGAPLPPAMQFPLKXDRISTFKNMPFLGCACTPERMAAGFIHCPTENEPDLAQCFFC 60

Db 1 MGAPLPPAMQFPLKXDRISTFKNMPFLGCACTPERMAAGFIHCPTENEPDLAQCFFC 60

Qy 61 FKELEGEPPDDPIEHHKHSSGCAFLSVKQFEEITLGEFLKDRERAKNKIAKETNNK 120

Db 61 FKELEGEPPDDPIEHHKHSSGCAFLSVKQFEEITLGEFLKDRERAKNKIAKETNNK 120

Qy 121 KKEFEETAKKVRRAIEQLAAMD 142

Db 121 KKEFEETAKKVRRAIEQLAAMD 142

RESULT 3

US-09-283-144-15

Sequence 15, Application US/09283144

Patent No. 6346389

GENERAL INFORMATION:

APPLICANT: Yale University

TITLE OF INVENTION: Method for Selectively Modulating the Interactions

FILE REFERENCE: 44574-5033-US

CURRENT APPLICATION NUMBER: US/09/283,144

CURRENT FILING DATE: 1999-04-01

EARLIER APPLICATION NUMBER: US 60/080,288

EARLIER FILING DATE: 1998-04-01

EARLIER APPLICATION NUMBER: US 08/975,080

EARLIER FILING DATE: 1997-11-20

EARLIER APPLICATION NUMBER: PCT/US97/21880

EARLIER FILING DATE: 1997-11-20

EARLIER APPLICATION NUMBER: US 60/031,435

EARLIER FILING DATE: 1996-11-20

NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 15

LENGTH: 37

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Amino acid sequence from beginning of coding region of

US-09-283-144-15

Query Match 26.1%; Score 37; DB 4; Length 37;

Best Local Similarity 100.0%; Pred. No. 5.2e-31; Indels 0; Gaps 0;

Matches 37; Conservative 0; Mismatches 0; Indels 0;

Qy 1 MGAPLPPAMQFPLKXDRISTFKNMPFLGCACTPER 37

Db 1 MGAPLPPAWQPLKDHRISTFKWMPFLGCACTPFR 37

## RESULT 4

US-08-975-080-4  
Sequence 4, Application US/08975080  
Patent No. 6245523  
GENERAL INFORMATION:  
APPLICANT: Altieri, Dario C.  
TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS  
CELLULAR APOPTOSIS, AND ITS MODULATION  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
STREET: 1800 M Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036-5869  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/975,080  
FILING DATE: 20-NOV-1997  
PRIOR APPLICATION NUMBER: US 60/031,435  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Reid G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 044574-5022-01-WO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-467-7176  
TELEFAX: 202-467-7000  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-975-080-4

Query Match 14.1%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1e-13;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 EGMEPDDPPIEBHKHSSGC 84  
Db 1 EGMEPDDPPIEBHKHSSGC 20

RESULT 5  
US-08-975-080-3  
Sequence 3, Application US/08975080  
Patent No. 6245523  
GENERAL INFORMATION:  
APPLICANT: Altieri, Dario C.  
TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS  
CELLULAR APOPTOSIS, AND ITS MODULATION  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
STREET: 1800 M Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036-5869  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/975,080  
FILING DATE: 20-NOV-1997  
PRIOR APPLICATION NUMBER: US 60/031,435  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Reid G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 044574-5022-01-WO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-467-7176  
TELEFAX: 202-467-7000  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-975-080-3

Query Match 12.0%; Score 17; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.1e-10;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPPAWQPLKDHRI 19  
Db 1 APTLPPAWQPLKDHRI 17

RESULT 6  
US-09-515-514-1  
Sequence 1, Application US/09515514  
Patent No. 6509162  
GENERAL INFORMATION:  
APPLICANT: Altieri, Dario C.  
TITLE OF INVENTION: Methods for Selectively Modulating Survivin Apoptosis  
PATHWAYS  
FILE REFERENCE: 44574-5063-US  
CURRENT APPLICATION NUMBER: US/09/515,514  
CURRENT FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 1  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (7)  
OTHER INFORMATION: PHOSPHORYLATION  
US-09-515-514-1

Query Match 12.0%; Score 17; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.1e-10;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LBGACTPERMAAGPI 44  
Db 1 LBGACTPERMAAGPI 17

RESULT 7  
US-09-515-514-12  
Sequence 12, Application US/09515514  
Patent No. 6509162  
GENERAL INFORMATION:  
APPLICANT: Altieri, Dario C.

```

; TITLE OF INVENTION: Methods for Selectively Modulating Survivin Apoptosis
; FILE REFERENCE: 44574-5063-US
; CURRENT APPLICATION NUMBER: US/09/515,514
; CURRENT FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Survivin phosphorylation motif
US-09-515-514-12

Query Match      6.3%; Score 9; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Gaps 0;

QY      31 CACTPERMA 39
Db      1 CACTPERMA 9

RESULT 8
US-09-515-514-13
; Sequence 13, Application US/09515514
; Patent No. 6509162
; GENERAL INFORMATION:
; APPLICANT: Altieri, Dario C.
; TITLE OF INVENTION: Methods for Selectively Modulating Survivin Apoptosis
; FILE REFERENCE: 44574-5063-US
; CURRENT APPLICATION NUMBER: US/09/515,514
; CURRENT FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: Survivin phosphorylation motif
US-09-515-514-13

Query Match      6.3%; Score 9; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      31 CACTPERMA 39
Db      1 CACTPERMA 9

RESULT 9
US-09-205-258-450
; Sequence 450, Application US/09205258
; Patent No. 6528174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; SOFTWARE: PatentIn Ver. 2.0
```

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
```

SEQ ID NO 450  
LENGTH: 101  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (44)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (46)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (77)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (78)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (101)  
OTHER INFORMATION: Xaa equals stop translation  
US-09-205-258-450

Query Match 4.9%; Score 7; DB 4; Length 101;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 13 FLKDHRI 19  
Db 22 FLKDHRI 28

RESULT 10  
US-09-732-210-281  
Sequence 281, Application US/097322210  
Patent No. 6573361  
GENERAL INFORMATION:  
APPLICANT: Bunkers, Greg J.  
APPLICANT: Liang, Jihong  
APPLICANT: Miltanck, Cindy A.  
APPLICANT: Seale, Jeffrey W.  
APPLICANT: Wu, Yonnie S.  
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
FILE REFERENCE: 38-21150361B  
CURRENT APPLICATION NUMBER: US/09/732,210  
CURRENT FILING DATE: 2000-12-07  
PRIOR APPLICATION NUMBER: US 60/169,513  
PRIOR FILING DATE: 1999-12-07  
PRIOR APPLICATION NUMBER: US 60/169,340  
PRIOR FILING DATE: 1999-12-07  
NUMBER OF SEQ ID NOS: 1753  
SEQ ID NO 281  
LENGTH: 130  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-09-732-210-281

Query Match 4.9%; Score 7; DB 4; Length 130;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 94 EELTIGE 100  
Db 47 EELTIGE 53

RESULT 11  
US-09-107-532A-5332  
Sequence 5332, Application US/09107532A  
Patent No. 6583275

GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSER: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Denake  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-8277  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5332:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 167 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (8) LOCATION 1...167  
SEQUENCE DESCRIPTION: SEQ ID NO: 5332:  
US-09-107-532A-5332

Query Match 4.9%; Score 7; DB 4; Length 167;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 131 VRRARQ 137  
Db 145 VRRARQ 151

RESULT 12  
US-09-107-532A-5027  
Sequence 5027, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSER: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneka  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5027:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 186 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (B) LOCATION 1..186  
SEQUENCE DESCRIPTION: SEQ ID NO: 5027:  
US-09-107-532A-5027

Query Match 4.9% Score 7; DB 4; Length 186;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 TAKVRR 133  
DB 43 TAKVRR 49

RESULT 13  
US-09-134-000C-4317  
Sequence 4317, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4317  
LENGTH: 208  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-4317

Query Match 4.9% Score 7; DB 4; Length 208;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 TAKVRR 133  
DB 65 TAKVRR 71

RESULT 14  
US-09-134-000C-3780  
Sequence 3780, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3780  
LENGTH: 258  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-3780

Query Match 4.9% Score 7; DB 4; Length 258;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 EETAKKV 131  
DB 171 EETAKKV 177

RESULT 15  
US-09-475-316A-70  
Sequence 70, Application US/09475316A  
Patent No. 6210942  
GENERAL INFORMATION:  
APPLICANT: Lewis, No. 6210942man G.  
APPLICANT: Dinkova-Kostova, Albena T.  
APPLICANT: Fujita, Masayuki  
APPLICANT: Gang, David R.  
APPLICANT: Sakkenen, Simo  
APPLICANT: Ford, Joshua D  
TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASES,  
FILE REFERENCE: WSUR-1-13793  
CURRENT APPLICATION NUMBER: US/09/475,316A  
CURRENT FILING DATE: 1999-12-30  
PRIOR APPLICATION NUMBER: 09/307,653  
PRIOR FILING DATE: 1999-05-07  
PRIOR APPLICATION NUMBER: PCT/US97/20391  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: 60/054,380  
PRIOR FILING DATE: 1997-07-31  
PRIOR APPLICATION NUMBER: 60/030,522  
PRIOR FILING DATE: 1996-11-08  
NUMBER OF SEQ ID NOS: 122  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 70  
LENGTH: 307  
TYPE: PRT  
ORGANISM: Tsuga heterophylla  
US-09-475-316A-70

Query Match 4.9% Score 7; DB 3; Length 307;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAIE 136  
DB 135 KVRRAIE 141

RESULT 16



US-09-704-640-70  
; Sequence 70, Application US/09704640  
; Patent No. 6635459  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, No. 6635459man G.  
; APPLICANT: Davin, Laurence B.  
; APPLICANT: Dinkova-Kostova, Albena T.  
; APPLICANT: Fujita, Masayuki  
; APPLICANT: Gang, David R.  
; APPLICANT: Sarkanen, Simo  
; APPLICANT: Ford, Joshua D  
; TITLE OF INVENTION: RECOMBINANT PINOSESINOL/LARICRESINOL REDUCTASE,  
; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEIN AND METHODS OF USE  
; FILE REFERENCE: MSUR-1-16492  
; CURRENT APPLICATION NUMBER: US/09/704,640  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: 09/475,316  
; PRIOR FILING DATE: 1999-12-30  
; PRIOR APPLICATION NUMBER: 09/307,653  
; PRIOR FILING DATE: 1999-05-07  
; PRIOR APPLICATION NUMBER: PCT/US97/20391  
; PRIOR FILING DATE: 1997-11-07  
; PRIOR APPLICATION NUMBER: 60/054,380  
; PRIOR FILING DATE: 1997-07-31  
; PRIOR APPLICATION NUMBER: 60/030,522  
; PRIOR FILING DATE: 1996-11-08  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 70  
; LENGTH: 307  
; TYPE: PRT  
; ORGANISM: Tsuga heterophylla  
US-09-704-640-70

Query Match 4.9%; Score 7; DB 4; Length 307;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAIE 136  
Db 135 KVRRAIE 141

RESULT 17  
US-09-475-316A-72  
; Sequence 72, Application US/09475316A  
; Patent No. 6210942  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, No. 6210942man G.  
; APPLICANT: Davin, Laurence B.  
; APPLICANT: Dinkova-Kostova, Albena T.  
; APPLICANT: Fujita, Masayuki  
; APPLICANT: Gang, David R.  
; APPLICANT: Sarkanen, Simo  
; APPLICANT: Ford, Joshua D  
; TITLE OF INVENTION: RECOMBINANT PINOSESINOL/LARICRESINOL REDUCTASES,  
; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE  
; FILE REFERENCE: MSUR-1-13793  
; CURRENT APPLICATION NUMBER: US/09/475,316A  
; PRIOR FILING DATE: 1999-12-30  
; PRIOR APPLICATION NUMBER: 09/307,653  
; PRIOR FILING DATE: 1999-05-07  
; PRIOR APPLICATION NUMBER: PCT/US97/20391  
; PRIOR FILING DATE: 1997-11-07  
; PRIOR APPLICATION NUMBER: 60/054,380  
; PRIOR FILING DATE: 1997-07-31  
; PRIOR APPLICATION NUMBER: 60/030,522  
; PRIOR FILING DATE: 1996-11-08  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 72  
; LENGTH: 309  
; TYPE: PRT

US-09-475-316A-72  
; Sequence 72, Application US/09475316A  
; Patent No. 6210942  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, No. 6210942man G.  
; APPLICANT: Davin, Laurence B.  
; APPLICANT: Dinkova-Kostova, Albena T.  
; APPLICANT: Fujita, Masayuki  
; APPLICANT: Gang, David R.  
; APPLICANT: Sarkanen, Simo  
; APPLICANT: Ford, Joshua D  
; TITLE OF INVENTION: RECOMBINANT PINOSESINOL/LARICRESINOL REDUCTASES,  
; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE  
; FILE REFERENCE: MSUR-1-13793  
; CURRENT APPLICATION NUMBER: US/09/475,316A  
; PRIOR FILING DATE: 1999-12-30  
; PRIOR APPLICATION NUMBER: 09/307,653  
; PRIOR FILING DATE: 1999-05-07  
; PRIOR APPLICATION NUMBER: PCT/US97/20391  
; PRIOR FILING DATE: 1997-11-07  
; PRIOR APPLICATION NUMBER: 60/054,380  
; PRIOR FILING DATE: 1997-07-31  
; PRIOR APPLICATION NUMBER: 60/030,522  
; PRIOR FILING DATE: 1996-11-08  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 72  
; LENGTH: 309  
; TYPE: PRT

ORGANISM: Tsuga heterophylla  
US-09-475-316A-72

Query Match 4.9%; Score 7; DB 3; Length 309;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAIE 136  
Db 137 KVRRAIE 143

RESULT 18  
US-09-704-640-72  
; Sequence 72, Application US/09704640  
; Patent No. 6635459  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, No. 6635459man G.  
; APPLICANT: Davin, Laurence B.  
; APPLICANT: Dinkova-Kostova, Albena T.  
; APPLICANT: Fujita, Masayuki  
; APPLICANT: Gang, David R.  
; APPLICANT: Sarkanen, Simo  
; APPLICANT: Ford, Joshua D  
; TITLE OF INVENTION: RECOMBINANT PINOSESINOL/LARICRESINOL REDUCTASE,  
; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEIN AND METHODS OF USE  
; FILE REFERENCE: MSUR-1-16492  
; CURRENT APPLICATION NUMBER: US/09/704,640  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: 09/475,316  
; PRIOR FILING DATE: 1999-12-30  
; PRIOR APPLICATION NUMBER: 09/307,653  
; PRIOR FILING DATE: 1999-05-07  
; PRIOR APPLICATION NUMBER: PCT/US97/20391  
; PRIOR FILING DATE: 1997-11-07  
; PRIOR APPLICATION NUMBER: 60/054,380  
; PRIOR FILING DATE: 1997-07-31  
; PRIOR APPLICATION NUMBER: 60/030,522  
; PRIOR FILING DATE: 1996-11-08  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 72  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Tsuga heterophylla  
US-09-704-640-72

Query Match 4.9%; Score 7; DB 4; Length 309;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAIE 136  
Db 137 KVRRAIE 143

RESULT 19  
US-09-475-316A-62  
; Sequence 62, Application US/09475316A  
; Patent No. 6210942  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, No. 6210942man G.  
; APPLICANT: Davin, Laurence B.  
; APPLICANT: Dinkova-Kostova, Albena T.  
; APPLICANT: Fujita, Masayuki  
; APPLICANT: Gang, David R.  
; APPLICANT: Sarkanen, Simo  
; APPLICANT: Ford, Joshua D  
; TITLE OF INVENTION: RECOMBINANT PINOSESINOL/LARICRESINOL REDUCTASES,  
; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE  
; FILE REFERENCE: MSUR-1-13793  
; CURRENT APPLICATION NUMBER: US/09/475,316A  
; PRIOR FILING DATE: 1999-12-30

US-09-475-316A-62  
; Sequence 62, Application US/09475316A  
; Patent No. 6210942  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, No. 6210942man G.  
; APPLICANT: Davin, Laurence B.  
; APPLICANT: Dinkova-Kostova, Albena T.  
; APPLICANT: Fujita, Masayuki  
; APPLICANT: Gang, David R.  
; APPLICANT: Sarkanen, Simo  
; APPLICANT: Ford, Joshua D  
; TITLE OF INVENTION: RECOMBINANT PINOSESINOL/LARICRESINOL REDUCTASES,  
; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE  
; FILE REFERENCE: MSUR-1-13793  
; CURRENT APPLICATION NUMBER: US/09/475,316A  
; PRIOR FILING DATE: 1999-12-30

;; PRIOR APPLICATION NUMBER: 09/307,653  
;; PRIOR FILING DATE: 1999-05-07  
;; PRIOR APPLICATION NUMBER: PCT/US97/20391  
;; PRIOR FILING DATE: 1997-11-07  
;; PRIOR APPLICATION NUMBER: 60/054,380  
;; PRIOR FILING DATE: 1997-07-31  
;; PRIOR APPLICATION NUMBER: 60/030,522  
;; PRIOR FILING DATE: 1996-11-08  
;; NUMBER OF SEQ ID NOS: 122  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO: 62  
;; LENGTH: 313  
;; TYPE: PRT  
;; ORGANISM: Thuja plicata  
US-09-475-316A-62

Query Match 4.9%; Score 7; DB 3; Length 313;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 KVRRAIE 136  
Db 140 KVRRAIE 146

RESULT 20  
US-09-704-640-62  
;; Sequence 62, Application US/09704640  
;; Patent No. 6635459  
;; GENERAL INFORMATION:  
;; APPLICANT: Lewis, No. 6635459man G.  
;; APPLICANT: Davin, Laurence B, Albena T.  
;; APPLICANT: Dinkova-Kostova, Albena T.  
;; APPLICANT: Fujita, Masayuki  
;; APPLICANT: Gang, David R.  
;; APPLICANT: Sarkanen, Simo  
;; APPLICANT: Ford, Joshua D  
;; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASE,  
;; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEIN AND METHODS OF USE  
;; FILE REFERENCE: MSUR-1-16492  
;; CURRENT APPLICATION NUMBER: US/09/704,640  
;; PRIOR FILING DATE: 2000-11-02  
;; PRIOR APPLICATION NUMBER: 09/475,316  
;; PRIOR FILING DATE: 1999-12-30  
;; PRIOR APPLICATION NUMBER: 09/307,653  
;; PRIOR FILING DATE: 1999-05-07  
;; PRIOR APPLICATION NUMBER: PCT/US97/20391  
;; PRIOR FILING DATE: 1997-11-07  
;; PRIOR APPLICATION NUMBER: 60/054,380  
;; PRIOR FILING DATE: 1997-07-31  
;; PRIOR APPLICATION NUMBER: 60/030,522  
;; PRIOR FILING DATE: 1996-11-08  
;; NUMBER OF SEQ ID NOS: 122  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO: 62  
;; LENGTH: 313  
;; TYPE: PRT  
;; ORGANISM: Thuja plicata  
US-09-704-640-62

Query Match 4.9%; Score 7; DB 4; Length 313;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 KVRRAIE 136  
Db 140 KVRRAIE 146

RESULT 21  
US-09-475-316A-66  
;; Sequence 66, Application US/09475316A  
;; Patent No. 6210942

;; GENERAL INFORMATION:  
;; APPLICANT: Lewis, No. 6210942man G.  
;; APPLICANT: Davin, Laurence B, Albena T.  
;; APPLICANT: Dinkova-Kostova, Albena T.  
;; APPLICANT: Fujita, Masayuki  
;; APPLICANT: Gang, David R.  
;; APPLICANT: Sarkanen, Simo  
;; APPLICANT: Ford, Joshua D  
;; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASES,  
;; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE  
;; FILE REFERENCE: MSUR-1-13793  
;; CURRENT APPLICATION NUMBER: US/09/475,316A  
;; PRIOR FILING DATE: 1999-12-30  
;; PRIOR APPLICATION NUMBER: 09/307,653  
;; PRIOR FILING DATE: 1999-05-07  
;; PRIOR APPLICATION NUMBER: PCT/US97/20391  
;; PRIOR FILING DATE: 1997-11-07  
;; PRIOR APPLICATION NUMBER: 60/054,380  
;; PRIOR FILING DATE: 1997-07-31  
;; PRIOR APPLICATION NUMBER: 60/030,522  
;; PRIOR FILING DATE: 1996-11-08  
;; NUMBER OF SEQ ID NOS: 122  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO: 66  
;; LENGTH: 314  
;; TYPE: PRT  
;; ORGANISM: Thuja plicata  
US-09-475-316A-66

Query Match 4.9%; Score 7; DB 3; Length 314;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 KVRRAIE 136  
Db 140 KVRRAIE 146

RESULT 22  
US-09-475-316A-118  
;; Sequence 118, Application US/09475316A  
;; Patent No. 6210942  
;; GENERAL INFORMATION:  
;; APPLICANT: Lewis, No. 6210942man G.  
;; APPLICANT: Davin, Laurence B, Albena T.  
;; APPLICANT: Dinkova-Kostova, Albena T.  
;; APPLICANT: Fujita, Masayuki  
;; APPLICANT: Gang, David R.  
;; APPLICANT: Sarkanen, Simo  
;; APPLICANT: Ford, Joshua D  
;; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASES,  
;; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE  
;; FILE REFERENCE: MSUR-1-13793  
;; CURRENT APPLICATION NUMBER: US/09/475,316A  
;; PRIOR FILING DATE: 1999-12-30  
;; PRIOR APPLICATION NUMBER: 09/307,653  
;; PRIOR FILING DATE: 1999-05-07  
;; PRIOR APPLICATION NUMBER: PCT/US97/20391  
;; PRIOR FILING DATE: 1997-11-07  
;; PRIOR APPLICATION NUMBER: 60/054,380  
;; PRIOR FILING DATE: 1997-07-31  
;; PRIOR APPLICATION NUMBER: 60/030,522  
;; PRIOR FILING DATE: 1996-11-08  
;; NUMBER OF SEQ ID NOS: 122  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO: 118  
;; LENGTH: 314  
;; TYPE: PRT  
;; ORGANISM: Schisandra chinensis  
US-09-475-316A-118

Query Match 4.9%; Score 7; DB 3; Length 314;  
Best Local Similarity 100.0%; Pred. No. 26;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAE 136  
 |||||  
 Db 142 KVRRAE 148

## RESULT 23

US-09-704-640-66  
 ; Sequence 66, Application US/09704640  
 ; Patent No. 6635459  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lewis, No. 6635459man G.  
 ; APPLICANT: Davin, Laurence B.  
 ; APPLICANT: Dinkova-Kostova, Albena T.  
 ; APPLICANT: Fujita, Masayuki  
 ; APPLICANT: Gang, David R.  
 ; APPLICANT: Sarkanen, Simo  
 ; APPLICANT: Ford, Joshua D  
 ; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LAICIREBINOL REDUCTASE,  
 ; FILE REFERENCE: MSUR-1-16492  
 ; CURRENT APPLICATION NUMBER: US/09/704,640  
 ; PRIOR FILING DATE: 2000-11-02  
 ; PRIOR APPLICATION NUMBER: 09/475,316  
 ; PRIOR FILING DATE: 1999-12-30  
 ; PRIOR APPLICATION NUMBER: 09/307,653  
 ; PRIOR FILING DATE: 1999-05-07  
 ; PRIOR APPLICATION NUMBER: PCT/US97/20391  
 ; PRIOR FILING DATE: 1997-11-07  
 ; PRIOR APPLICATION NUMBER: 60/054,380  
 ; PRIOR FILING DATE: 1997-07-31  
 ; PRIOR APPLICATION NUMBER: 60/030,522  
 ; PRIOR FILING DATE: 1996-11-08  
 ; NUMBER OF SEQ ID NOS: 122  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 66  
 ; LENGTH: 314  
 ; TYPE: PRT  
 ; ORGANISM: Thruja plicata  
 ; US-09-704-640-66

Query Match 4.9%; Score 7; DB 4; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAE 136  
 |||||  
 Db 140 KVRRAE 146

## RESULT 24

US-09-704-640-118  
 ; Sequence 118, Application US/09704640  
 ; Patent No. 6635459  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lewis, No. 6635459man G.  
 ; APPLICANT: Davin, Laurence B.  
 ; APPLICANT: Dinkova-Kostova, Albena T.  
 ; APPLICANT: Fujita, Masayuki  
 ; APPLICANT: Gang, David R.  
 ; APPLICANT: Sarkanen, Simo  
 ; APPLICANT: Ford, Joshua D  
 ; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LAICIREBINOL REDUCTASE,  
 ; FILE REFERENCE: MSUR-1-16492  
 ; CURRENT APPLICATION NUMBER: US/09/704,640  
 ; PRIOR FILING DATE: 2000-11-02  
 ; PRIOR APPLICATION NUMBER: 09/475,316  
 ; PRIOR FILING DATE: 1999-12-30  
 ; PRIOR APPLICATION NUMBER: 09/307,653  
 ; PRIOR FILING DATE: 1999-05-07  
 ; PRIOR APPLICATION NUMBER: PCT/US97/20391

;; PRIOR FILING DATE: 1997-11-07  
 ; PRIOR APPLICATION NUMBER: 60/054,380  
 ; PRIOR FILING DATE: 1997-07-31  
 ; PRIOR APPLICATION NUMBER: 60/030,522  
 ; PRIOR FILING DATE: 1996-11-08  
 ; NUMBER OF SEQ ID NOS: 122  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 118  
 ; LENGTH: 314  
 ; TYPE: PRT  
 ; ORGANISM: Schisandra chinensis  
 ; US-09-704-640-118

Query Match 4.9%; Score 7; DB 4; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KVRRAE 136  
 |||||  
 Db 142 KVRRAE 148

RESULT 25  
 US-09-252-991A-25075  
 ; Sequence 25075, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 25075  
 ; LENGTH: 418  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-25075

Query Match 4.9%; Score 7; DB 4; Length 418;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PAMQPL 14  
 |||||  
 Db 406 PAMQPL 412

RESULT 26  
 US-09-107-532A-5734  
 ; Sequence 5734, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street  
 ; CITY: Waltham  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02354  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD-ROM ISO9660  
 ; COMPUTER: PC  
 ; OPERATING SYSTEM: <Unknown>

```
/
/   SOFTWARE: ASCII
/   CURRENT APPLICATION DATA:
/   APPLICATION NUMBER: US/09/107,532A
/   FILING DATE: 30-Jun-1998
/   PRIOR APPLICATION DATA:
/   APPLICATION NUMBER: 60/085,598
/   FILING DATE: 14 May 1998
/   APPLICATION NUMBER: 60/051571
/   FILING DATE: July 2, 1997
/   ATTORNEY/AGENT INFORMATION:
/   NAME: Ariniello, Pamela Deneke
/   REGISTRATION NUMBER: 40,489
/   REFERENCE/DOCKET NUMBER: GTC-012
/   TELECOMMUNICATION INFORMATION:
/   TELEPHONE: (781)893-5007
/   TELEFAX: (781)893-8277
/   INFORMATION FOR SEQ ID NO: 5734:
/   SEQUENCE CHARACTERISTICS:
/   LENGTH: 580 amino acids
/   TYPE: amino acid
/   TOPOLOGY: linear
/   MOLECULE TYPE: protein
/   HYPOTHEICAL: YES
/   ORIGINAL SOURCE:
/   ORGANISM: Enterococcus faecium
/   FEATURE:
/   NAME/KEY: misc feature
/   LOCATION: (B) LOCATION 1...580
/   SEQUENCE DESCRIPTION: SEQ ID NO: 5734:
US-09-107-532A-5734

Query Match
Best Local Similarity 4.9%; Score 7; DB 4; Length 580;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 NKIAKET 117
DB 227 NKIAKET 233

RESULT 27
US-09-308-375-2
/ Sequence 2, Application US/09308375
/ Patent No. 6300117
/ GENERAL INFORMATION:
/ APPLICANT: Genencor International, Inc.
/ TITLE OF INVENTION: Proteases From Gram-Positive Organisms
/ FILE REFERENCE: GC394-PCT
/ CURRENT APPLICATION NUMBER: US/09/308,375
/ CURRENT FILING DATE: 1999-05-14
/ EARLIER APPLICATION NUMBER: EP9719636.4
/ EARLIER FILING DATE: 1997-03-15
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2
/ LENGTH: 2285
/ TYPE: PRT
/ ORGANISM: Bacillus subtilis
US-09-308-375-2

Query Match
Best Local Similarity 4.9%; Score 7; DB 4; Length 2285;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 RRAIFOL 138
DB 184 RRAIFOL 190

RESULT 28
US-08-928-383B-5
/ Sequence 5, Application US/08928383B
/ Patent No. 6210921
```

```
/
/   GENERAL INFORMATION:
/   APPLICANT: Robert W. Finberg, Jeffrey M. Bergelson,
/   APPLICANT: and Marshall S. Horwitz
/   TITLE OF INVENTION: CAR, A No. 6210921el Cocksackievirus and Adenovirus
/   TITLE OF INVENTION: Receptor
/   NUMBER OF SEQUENCES: 26
/   CORRESPONDENCE ADDRESS:
/   ADDRESSEE: LAHIVE & COCKFIELD, LLP
/   STREET: 28 State Street
/   CITY: Boston
/   STATE: Massachusetts
/   COUNTRY: USA
/   ZIP: 02109
/   COMPUTER READABLE FORM:
/   MEDIUM TYPE: floppy disk
/   COMPUTER: IBM PC compatible
/   OPERATING SYSTEM: PC-DOS/MS-DOS
/   SOFTWARE: Patent Release #1.0, Version #1.25
/   CURRENT APPLICATION DATA:
/   APPLICATION NUMBER: US/08/928,383B
/   FILING DATE: 12-SEP-1997
/   CLASSIFICATION:
/   PRIOR APPLICATION DATA:
/   APPLICATION NUMBER: US 60/026,100
/   FILING DATE: 13-SEP-1996
/   ATTORNEY/AGENT INFORMATION:
/   NAME: Mandragouras, Amy E.
/   REGISTRATION NUMBER: 36,207
/   REFERENCE/DOCKET NUMBER: DEN-020
/   TELECOMMUNICATION INFORMATION:
/   TELEPHONE: (617)227-7400
/   TELEFAX: (617)742-4214
/   INFORMATION FOR SEQ ID NO: 5:
/   SEQUENCE CHARACTERISTICS:
/   LENGTH: 11 amino acids
/   TYPE: amino acid
/   TOPOLOGY: linear
/   MOLECULE TYPE: peptide
/   FRAGMENT TYPE: internal
US-08-928-383B-5

Query Match
Best Local Similarity 4.2%; Score 6; DB 3; Length 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PTLPPA 9
DB 5 PTLPPA 10

RESULT 29
US-08-764-640-130
/ Sequence 130, Application US/08764640
/ Patent No. 5869451
/ Patent No. 5869451 5837683
/ GENERAL INFORMATION:
/ APPLICANT: Dower, William J.
/ APPLICANT: Bartlett, Ronald W.
/ APPLICANT: Cwila, Steven E.
/ APPLICANT: Gates, Christian
/ APPLICANT: Schatz, Peter J.
/ APPLICANT: Balasubramanian, Palaniappan
/ APPLICANT: Mastrom, Christopher R.
/ APPLICANT: Hendren, Richard W.
/ APPLICANT: Depirnce, Randolph B.
/ APPLICANT: Poddaturi, Surekha
/ APPLICANT: Yin, Qun
/ TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
/ TITLE OF INVENTION: RECEPTOR
/ NUMBER OF SEQUENCES: 244
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Glaxo Wellcome
/ STREET: Five Moore Drive, P.O. Box 13398
```

CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/764,640  
FILING DATE: 11-DEC-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3281  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 130:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-764-640-130

Query Match 4.2%; Score 6; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 TLGSFL 102  
Db 2 TLGSFL 7

RESULT 30  
US-08-973-225-130  
Sequence 130, Application US/08973225A  
Patent No. 6083913  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
Barrett, Ronald W.  
Cwirla, Steven E.  
Duffin, David J.  
Gates, Christian  
Haselden, Sherill S.  
Matheakis, Larry C.  
Schatz, Peter J.  
Wagstrom, Christopher R.  
Wrighton, Nicholas C.  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
THROMBOPOIETIN RECEPTOR  
NUMBER OF SEQUENCES: 232  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/973,225A  
FILING DATE: 04-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392

REFERENCE/DOCKET NUMBER: PK3065USW  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 130:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 130:  
US-08-973-225-130

Query Match 4.2%; Score 6; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 TLGSFL 102  
Db 2 TLGSFL 7

RESULT 31  
US-09-244-298A-130  
Sequence 130, Application US/09244298A  
Patent No. 6121238  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
Barrett, Ronald W.  
Cwirla, Steven E.  
Gates, Christian  
Schatz, Peter J.  
Balasubramanian, Palaniappan  
Wagstrom, Christopher R.  
Hendren, Richard W.  
Depina, Randolph B.  
Podduri, Surekha  
APPLICANT: Yin, Qun  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
RECEPTOR  
NUMBER OF SEQUENCES: 244  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/244,298A  
FILING DATE: 11-DEC-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3281  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 130:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-244-298A-130

Query Match 4.2%; Score 6; DB 3; Length 12;

Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 TLGEFL 102  
Db 2 TLGEFL 7

RESULT 32  
US-09-516-704-130  
; Sequence 130, Application US/09516704  
; Patent No. 6251864

GENERAL INFORMATION:

APPLICANT: Dower, William J.  
Barrett, Ronald W.  
Cwirla, Steven E.  
Gates, Christian  
Schatz, Peter J.  
Balasubramanian, Palaniappan  
Wagstrom, Christopher R.  
Hendren, Richard W.  
Depinche, Randolph B.  
Podduri, Surekha  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A RECEPTOR

NUMBER OF SEQUENCES: 244

CORRESPONDENCE ADDRESS:

ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/09/516,704

APPLICATION NUMBER: 01-Mar-2000

FILING DATE: 01-Mar-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3261

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-248-1000

INFORMATION FOR SEQ ID NO: 130:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 130:

US-09-516-704-130

Query Match 4.2%; Score 6; DB 3; Length 12;

Best Local Similarity 100.0%; Pred. No. 17;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 TLGEFL 102  
Db 2 TLGEFL 7

RESULT 33  
US-09-549-090-130  
; Sequence 130, Application US/09549090  
; Patent No. 6465430

GENERAL INFORMATION:  
APPLICANT: Dower, William J.

Barrett, Ronald W.  
Cwirla, Steven E.  
Duffin, David J.  
Gates, Christian  
Haselden, Sherill S.  
Matheakis, Larry C.  
Schatz, Peter J.  
Wagstrom, Christopher R.  
Wrighton, Nicholas C.

TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A THROMBOPOIETIN RECEPTOR

NUMBER OF SEQUENCES: 232

CORRESPONDENCE ADDRESS:

ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/09/549,090

APPLICATION NUMBER: 13-Apr-2000

FILING DATE: 13-Apr-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/973,225

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3065USW

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-248-1000

INFORMATION FOR SEQ ID NO: 130:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 130:

US-09-549-090-130

Query Match 4.2%; Score 6; DB 4; Length 12;

Best Local Similarity 100.0%; Pred. No. 17;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 TLGEFL 102  
Db 2 TLGEFL 7

RESULT 34  
US-09-832-230A-130  
; Sequence 130, Application US/09832230A  
; Patent No. 6506362

GENERAL INFORMATION:

APPLICANT: Dower, William J et al

TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A RECEPTOR

NUMBER OF SEQUENCES: 244

CORRESPONDENCE ADDRESS:

ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/832,230A  
FILING DATE: 10-Apr-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3281  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 130:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 130:  
US-09-832-230A-130

Query Match  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 TLGEFL 102  
Db 2 TLGEFL 7

RESULT 35  
US-08-185-432-13  
Sequence 13, Application US/08185432  
Patent No. 5750652  
GENERAL INFORMATION:  
APPLICANT: Artavanis-Tsakonas, Spyridon  
APPLICANT: Bussseau, Isabelle  
APPLICANT: Diederich, Robert J.  
APPLICANT: Xu, Tian  
APPLICANT: Matsuno, Kenji  
TITLE OF INVENTION: DELTEx PROTEINS, NUCLEIC ACIDS, AND  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/185,432  
FILING DATE: 21-JAN-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 7326-006  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24 amino acids  
TYPE: amino acid

TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-185-432-13

Query Match  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLP 8  
Db 2 APTLP 7

RESULT 36  
US-08-936-165A-380  
Sequence 380, Application US/08936165A  
Patent No. 6348382  
GENERAL INFORMATION:  
APPLICANT: Black, Michael  
APPLICANT: Burnham, Martin  
APPLICANT: Hodgson, John  
APPLICANT: Knowles, David  
APPLICANT: Lonetto, Michael  
APPLICANT: Nicholas, Richard  
APPLICANT: Pratt, Julie  
APPLICANT: Reichard, Richard  
APPLICANT: Rosenberg, Martin  
APPLICANT: Ward, Judith  
TITLE OF INVENTION: No. 6348382el Prokaryotic polynucleotides,  
NUMBER OF SEQUENCES: 334  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: PASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/936,165A  
FILING DATE: 24-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/027,032  
FILING DATE: 24-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmal, Edward R.  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50549  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 380:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-936-165A-380

Query Match  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 AKNKA 114

Db 9 AKKIA 14

RESULT 37  
US-09-621-976-6992  
; Sequence 6992, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6992  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-621-976-6992

Query Match 4.2%; Score 6; DB 4; Length 58;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 OPLKD 16  
Db 21 OPLKD 26

RESULT 38  
US-09-328-352-4702  
; Sequence 4702, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 4702  
; LENGTH: 80  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-4702

Query Match 4.2%; Score 6; DB 4; Length 80;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 125 EETAKK 130  
Db 19 EETAKK 24

RESULT 39  
US-09-107-532A-3689  
; Sequence 3689, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A. Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts

COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD/ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 3689:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 93 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc-feature  
LOCATION: (B) LOCATION 1...93  
SEQUENCE DESCRIPTION: SEQ ID NO: 3689:  
US-09-107-532A-3689

Query Match 4.2%; Score 6; DB 4; Length 93;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 90 KKOFER 95  
Db 30 KKOFER 35

RESULT 40  
US-09-252-991A-19558  
; Sequence 19558, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196-136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 19558  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19558

Query Match 4.2%; Score 6; DB 4; Length 95;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 TNNKKX 122



Db 30 TNNKKK 35

RESULT 41  
US-09-134-001C-5653  
; Sequence 5653, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-09  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 5653  
; LENGTH: 97  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5653

Query Match 4.2%; Score 6; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 KXKXAK 115  
Db 91 KXKXAK 96

RESULT 42  
US-09-621-976-7607  
; Sequence 7607, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 7607  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-621-976-7607

Query Match 4.2%; Score 6; DB 4; Length 99;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 KXHSYG 83  
Db 19 KXHSYG 24

RESULT 43  
US-09-489-039A-12130  
; Sequence 12130, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 12130  
; LENGTH: 130  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-12130

Query Match 4.2%; Score 6; DB 4; Length 130;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 RATEQL 138  
Db 32 RATEQL 37

RESULT 44  
US-09-252-991A-17914  
; Sequence 17914, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 17914  
; LENGTH: 133  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17914

Query Match 4.2%; Score 6; DB 4; Length 133;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 KXRAAI 135  
Db 7 KXRAAI 12

RESULT 45  
US-09-800-729-141  
; Sequence 141, Application US/09800729  
; Patent No. 6605592  
; GENERAL INFORMATION:  
; APPLICANT: N. et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 141  
; LENGTH: 134  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-141

Query Match 4.2%; Score 6; DB 4; Length 134;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 AKVRR 133  
DB 123 AKVRR 128

RESULT 46  
US-09-328-352-4961  
; Sequence 4961, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; NUMBER OF SEQ ID NOS: 04  
; SEQ ID NO 4961  
; LENGTH: 135  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-4961

Query Match 4.2%; Score 6; DB 4; Length 135;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 KKVRR 134  
DB 60 KKVRR 65

RESULT 47  
US-09-732-210-545  
; Sequence 545, Application US/09732210  
; Patent No. 6573361  
; GENERAL INFORMATION:  
; APPLICANT: Bunkers, Greg J.  
; APPLICANT: Liang, Jihong  
; APPLICANT: Mitnick, Candy A.  
; APPLICANT: Seale, Jeffrey W.  
; APPLICANT: Wn, Yonnie S.  
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
; FILE REFERENCE: 38-21(15036)B  
; CURRENT APPLICATION NUMBER: US/09/732,210  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,513  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,340  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 1753  
; SEQ ID NO 545  
; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-09-732-210-545

Query Match 4.2%; Score 6; DB 4; Length 141;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 KXKIAK 115  
DB 94 KXKIAK 99

RESULT 48  
US-07-998-003A-36  
; Sequence 36, Application US/07998003A

Patent No. 5643781  
; GENERAL INFORMATION:  
; APPLICANT: Suzuki, Shintaro  
; TITLE OF INVENTION: Protocadherin Materials and Methods  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
; STREET: 20 South Clark Street  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/998,003A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5643781and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 30903  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/346-5750  
; TELEFAX: 312/984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 148 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-998-003A-36

Query Match 4.2%; Score 6; DB 1; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 KLDRR 108  
DB 107 KLDRR 112

RESULT 49  
US-08-453-274B-36  
; Sequence 36, Application US/08453274B  
; Patent No. 5663300  
; GENERAL INFORMATION:  
; APPLICANT: Suzuki, Shintaro  
; TITLE OF INVENTION: Protocadherin Materials and Methods  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/453,274B  
; FILING DATE: 30-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5663300and, Greta E.

REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 32660  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 148 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-453-274B-36

Query Match 4.2%; Score 6; DB 1; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 KLDRE 108  
Db 107 KLDRE 112

RESULT 50  
US-08-453-695A-36  
Sequence 36, Application US/08453695A  
Patent No. 5708143  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: Protocadherin Materials and Methods  
NUMBER OF SEQUENCES: 115  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
ADDRESS: Borun  
STREET: 233 South Wacker, 6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,695A  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5708143and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 32658  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 148 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-453-695A-36

Query Match 4.2%; Score 6; DB 1; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 KLDRE 108  
Db 107 KLDRE 112

RESULT 51  
US-08-468-347-22  
Sequence 22, Application US/08468347  
Patent No. 5783421  
GENERAL INFORMATION:  
APPLICANT: Zeelon, Elsha P.  
APPLICANT: Werber, Moshe M.  
APPLICANT: Levamou, Avigdor  
TITLE OF INVENTION: NOVEL POLYPEPTIDE HAVING FACTOR Xa  
TITLE OF INVENTION: INHIBITORY ACTIVITY  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,347  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/225,442  
FILING DATE: 08-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0317/43020-A/JFW/ENB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-977-9550  
TELEFAX: 212-664-0525  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 148 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..148  
US-08-468-347-22

Query Match 4.2%; Score 6; DB 1; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 RISTF 23  
Db 135 RISTF 140

RESULT 52  
US-08-268-161A-36  
Sequence 36, Application US/08268161A  
Patent No. 5798224  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: Protocadherin Materials and Methods  
NUMBER OF SEQUENCES: 115  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
ADDRESS: Borun  
US-08-268-161A-36

STREET: 233 South Wacker, 6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/268,161A  
APPLICATION NUMBER: US/08/268,161A  
FILING DATE: June 27, 1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Young J. Suh  
REGISTRATION NUMBER: P-41,337  
REFERENCE/DOCKET NUMBER: 27866/32149  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 148 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-268-161A-36

Query Match 4.2% Score 6; DB 1; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 KIDRRR 108  
Db 107 KIDRRR 112

RESULT 53  
US-08-467-389-22  
Sequence 22, Application US/08467389  
Patent No. 5824641  
GENERAL INFORMATION:  
APPLICANT: Zeelon, Elisha P.  
APPLICANT: Werber, Moshe M.  
APPLICANT: Levaton, Avigdor  
TITLE OF INVENTION: NOVEL POLYPEPTIDE HAVING FACTOR Xa  
TITLE OF INVENTION: INHIBITORY ACTIVITY  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,389  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/225,442  
FILING DATE: 08-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0317/43020-A/JPM/EAB

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-977-9550  
TELEFAX: 212-664-0525  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 148 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..148  
US-08-467-389-22

Query Match 4.2% Score 6; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 RISTFK 23  
Db 135 RISTFK 140

RESULT 54  
US-08-779-379-22  
Sequence 22, Application US/08779379  
Patent No. 5858970  
GENERAL INFORMATION:  
APPLICANT: Zeelon, Elisha P.  
APPLICANT: Werber, Moshe M.  
APPLICANT: Levaton, Avigdor  
TITLE OF INVENTION: NOVEL POLYPEPTIDE HAVING FACTOR Xa  
TITLE OF INVENTION: INHIBITORY ACTIVITY  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/779,379  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/225,442  
FILING DATE: 08-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0317/43020-A/JPM/EAB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-977-9550  
TELEFAX: 212-664-0525  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 148 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..148  
US-08-779-379-22

Query Match 4.2%; Score 6; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred.No.1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 RSTFK 23  
Db 135 RSTFK 140

RESULT 55  
US-08-469-219-22  
Sequence 22, Application US/08469219  
Patent No. 5863534  
GENERAL INFORMATION:  
APPLICANT: Zeelon, Elisha P.  
APPLICANT: Werber, Moshe M.  
APPLICANT: Levanon, Avigdor  
TITLE OF INVENTION: NOVEL POLYPEPTIDE HAVING FACTOR Xa  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,219  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/225,442  
FILING DATE: 08-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0317/43020-A/JPM/EA8  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-977-9550  
TELEFAX: 212-664-0525  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 148 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..148  
US-08-469-219-22

Query Match 4.2%; Score 6; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred.No.1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 RSTFK 23  
Db 135 RSTFK 140

RESULT 56  
US-08-453-702A-36  
Sequence 36, Application US/08453702A  
Patent No. 5891706  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: Protocadherin Materials and Methods  
NUMBER OF SEQUENCES: 115  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Marshall, O'Toole, Gerstein, Murray, &  
ADDRESS: Borun  
STREET: 233 South Wacker, 6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,702A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5891706and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 32657  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 148 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-453-702A-36

Query Match 4.2%; Score 6; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred.No.1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 KLDKRR 108  
Db 107 KLDKRR 112

RESULT 57  
US-09-228-152-22  
Sequence 22, Application US/09228152  
Patent No. 6211341  
GENERAL INFORMATION:  
APPLICANT: Zeelon, Elisha P.  
APPLICANT: Werber, Moshe M.  
APPLICANT: Levanon, Avigdor  
TITLE OF INVENTION: NOVEL POLYPEPTIDE HAVING FACTOR Xa INHIBITORY ACTIVITY  
FILE REFERENCE: 43020aYa  
CURRENT APPLICATION NUMBER: US/09/228,152  
CURRENT FILING DATE: 1999-01-11  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 22  
LENGTH: 148  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Amino acid

OTHER INFORMATION: sequence of clone pSP65-Xa1-4.  
US-09-228-152-22

Query Match 4.2%; Score 6; DB 3; Length 148;  
Best Local Similarity 100.0%; Pred.No.1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 R1STFK 23  
DB 135 R1STFK 140

RESULT 58  
US-09-099-639-36

Sequence 36, Application US/09099639  
Patent No. 6262237

GENERAL INFORMATION:

APPLICANT: Suzuki, Shintaro

TITLE OF INVENTION: Protocadherin Materials and Methods

NUMBER OF SEQUENCES: 115

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
STREET: 233 South Wacker, 6300 Sears Tower

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/099,639

FILING DATE: 18 JUN 1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/263,161

FILING DATE: 27 JUN 1994

ATTORNEY/AGENT INFORMATION:

NAME: Greta E. No. 6262237and

REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 27866/34703

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 148 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-099-639-36

Query Match 4.2%; Score 6; DB 3; Length 148;  
Best Local Similarity 100.0%; Pred.No.1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 KLDRR 108  
DB 107 KLDRR 112

RESULT 59  
PCT-US93-12588-36

Sequence 36, Application PC/TUS9312588

GENERAL INFORMATION:

APPLICANT: Suzuki, Shintaro

TITLE OF INVENTION: Protocadherin Materials and Methods

NUMBER OF SEQUENCES: 107

CORRESPONDENCE ADDRESS:

US-09-099-639-36

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
ADDRESSEE: Borun  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/12588

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/998,003

FILING DATE: 29 DEC 1992

ATTORNEY/AGENT INFORMATION:

NAME: Noland, Greta E.

REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 31811

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 148 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US93-12588-36

Query Match 4.2%; Score 6; DB 5; Length 148;  
Best Local Similarity 100.0%; Pred.No.1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 KLDRR 108  
DB 107 KLDRR 112

RESULT 60  
PCT-US95-08071-36

Sequence 36, Application PC/TUS9508071

GENERAL INFORMATION:

APPLICANT: Suzuki, Shintaro

TITLE OF INVENTION: Protocadherin Materials and Methods

NUMBER OF SEQUENCES: 115

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
STREET: 6300 Sears Tower, 233 S. Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/08071

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/12588

FILING DATE: 23 DEC 1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/998,003  
FILING DATE: 29 DEC 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Noland, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 32149  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 148 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-08071-36

Query Match 4.2%; Score 6; DB 5; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 KLDNR 108  
Db 107 KLDNR 112

RESULT 61  
US-08-226-264-26  
Sequence 26, Application US/08226264  
Patent No. 5801017  
GENERAL INFORMATION:  
APPLICANT: Weider, Moshe M.  
APPLICANT: Zeelon, Elisha P.  
APPLICANT: Levanon, Avigdor  
APPLICANT: Guy, Rachel  
APPLICANT: Goldusc, Arie  
APPLICANT: Rigbi, Meir  
APPLICANT: Panel, Amos  
APPLICANT: Fischer, Meir  
TITLE OF INVENTION: PRODUCTION OF RECOMBINANT FACTOR XA  
TITLE OF INVENTION: INHIBITORS OF LEECH HIRUDO MEDICINALIS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/226,264  
FILING DATE: 08-APR-94  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 4001-A/JPW/GJG  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
TELEX:  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 153 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown

TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-226-264-26

Query Match 4.2%; Score 6; DB 1; Length 153;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 RISTFK 23  
Db 140 RISTFK 145

RESULT 62  
US-09-738-946-4  
Sequence 4, Application US/09738946  
Patent No. 6579701  
GENERAL INFORMATION:  
APPLICANT: EXELIXIS, INC.  
TITLE OF INVENTION: DROSOPHILA HOMOLOGUES OF GENES AND PROTEINS IMPLICATED IN CANCER  
TITLE OF INVENTION: METHODS OF USE  
FILE REFERENCE: EX00-043C  
CURRENT APPLICATION NUMBER: US/09/738,946  
PRIOR FILING DATE: 2000-12-14  
PRIOR APPLICATION NUMBER: 60/170,832  
PRIOR FILING DATE: 1999-12-14  
PRIOR APPLICATION NUMBER: 60/170,838  
PRIOR FILING DATE: 1999-12-14  
PRIOR APPLICATION NUMBER: 60/178,580  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/185,879  
PRIOR FILING DATE: 2000-02-29  
PRIOR APPLICATION NUMBER: 60/185,880  
PRIOR FILING DATE: 2000-02-29  
PRIOR APPLICATION NUMBER: 60/186,150  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 60/189,701  
PRIOR FILING DATE: 2000-03-15  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Version 3.0  
SEQ ID NO: 4  
LENGTH: 153  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-738-946-4

Query Match 4.2%; Score 6; DB 4; Length 153;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 MAAEAF 43  
Db 51 MAAEAF 56

RESULT 63  
US-09-732-210-1644  
Sequence 1644, Application US/09732210  
Patent No. 6573361  
GENERAL INFORMATION:  
APPLICANT: Bunkers, Greg J.  
APPLICANT: Liang, Jihong  
APPLICANT: Mitnick, Cindy A.  
APPLICANT: Seale, Jeffrey W.  
APPLICANT: Wu, Yonnie S.  
TITLE OF INVENTION: Anti-Fungal Proteins and Methods for Their Use  
FILE REFERENCE: 38-21(15036)B  
CURRENT APPLICATION NUMBER: US/09/732,210  
CURRENT FILING DATE: 2000-12-07  
PRIOR APPLICATION NUMBER: US 60/169,513

PRIOR FILING DATE: 1999-12-07  
PRIOR APPLICATION NUMBER: US 60/169,340  
PRIOR FILING DATE: 1999-12-07  
NUMBER OF SEQ ID NOS: 1753  
SEQ ID NO 1644  
LENGTH: 155  
TYPE: PR  
ORGANISM: Thermus aquaticus (subsp. thermophilus)  
US-09-732-210-1644

Query Match 4.2%; Score 6; DB 4; Length 155;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 ERARA 41  
Db 141 ERARA 146

RESULT 64  
US-09-134-000C-4018  
Sequence 4018, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 4018  
LENGTH: 163  
TYPE: PR  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-4018

Query Match 4.2%; Score 6; DB 4; Length 163;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 126 ERAKV 131  
Db 84 ERAKV 89

RESULT 65  
US-08-809-267-5  
Sequence 5, Application US/08809267  
Patent No. 5861296  
GENERAL INFORMATION:  
APPLICANT: LENNOX, Tricia L.  
APPLICANT: SLATKO, Barton E.  
APPLICANT: SEARS, Lauren E.  
TITLE OF INVENTION: PURIFIED THERMOSTABLE INORGANIC  
TITLE OF INVENTION: PYROPHOSPHATASE OBTAINABLE FROM THERMOCOCCUS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NEW ENGLAND BIOLABS, INC.  
STREET: 32 TOZER ROAD  
CITY: BEVERLY  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 01915  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,267  
FILING DATE: 12-MAR-1997  
CLASSIFICATION: 433  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13662  
FILING DATE:  
APPLICATION NUMBER: US 08/329,721  
FILING DATE: 25-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: WILLIAMS, GREGORY D.  
REGISTRATION NUMBER: 30901  
REFERENCE/DOCKET NUMBER: NEB-105-PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508) 927-5054  
TELEFAX: (508) 927-1705  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 175 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-809-267-5

Query Match 4.2%; Score 6; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 ERAKN 112  
Db 170 ERAKN 175

RESULT 66  
PCT-US95-13662A-5  
Sequence 5, Application PC/TUS9513662A  
GENERAL INFORMATION:  
APPLICANT: LENNOX, Tricia L.  
APPLICANT: SLATKO, Barton E.  
APPLICANT: SEARS, Lauren E.  
TITLE OF INVENTION: PURIFIED THERMOSTABLE INORGANIC  
TITLE OF INVENTION: PYROPHOSPHATASE OBTAINABLE FROM THERMOCOCCUS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NEW ENGLAND BIOLABS, INC.  
STREET: 32 TOZER ROAD  
CITY: BEVERLY  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 01915  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13662A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/329,721  
FILING DATE: 25-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: WILLIAMS, GREGORY D.  
REGISTRATION NUMBER: 30901  
REFERENCE/DOCKET NUMBER: NEB-105-PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508) 927-5054  
TELEFAX: (508) 927-1705  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:



LENGTH: 175 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCT-US95-13662A-5

Query Match 4.2%; Score 6; DB 5; Length 175;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 ERAKX 112  
Db 170 ERAKX 175

RESULT 67  
US-09-252-991A-16898  
Sequence 16898, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 16898  
LENGTH: 181  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16898

Query Match 4.2%; Score 6; DB 4; Length 181;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTLPP 8  
Db 130 APTLPP 135

RESULT 68  
US-09-328-352-4216  
Sequence 4216, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Berton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328.352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 4216  
LENGTH: 182  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-4216

Query Match 4.2%; Score 6; DB 4; Length 182;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 EULIGE 100  
Db 50 EULIGE 55

RESULT 69  
US-09-107-532A-6802  
Sequence 6802, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESS: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA

ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107.532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Denake  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 6802:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 194 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...194  
SEQUENCE DESCRIPTION: SEQ ID NO: 6802:  
US-09-107-532A-6802

Query Match 4.2%; Score 6; DB 4; Length 194;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 ETAKXV 131  
Db 119 ETAKXV 124

RESULT 70  
US-09-252-991A-16813  
Sequence 16813, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 16813  
; LENGTH: 204  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16813

Query Match  
Best Local Similarity 4.2%; Score 6; DB 4; Length 204;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PTLPPA 9  
Db 80 PTLPPA 85

RESULT 71  
US-09-328-352-6654  
; Sequence 6654, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; PRIOR FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6654  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6654

Query Match  
Best Local Similarity 4.2%; Score 6; DB 4; Length 207;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 KKKKFE 125  
Db 7 KKKKFE 12

RESULT 72  
US-09-252-991A-32926  
; Sequence 32926, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 32920  
; LENGTH: 248  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
; NAME/KEY: UNSURE  
; LOCATION: (216)  
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.  
US-09-252-991A-32920

Query Match 4.2%; Score 6; DB 4; Length 248;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 RMAEAG 42  
Db 147 RMAEAG 152

RESULT 73  
US-09-252-991A-32676  
; Sequence 32676, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 32676  
; LENGTH: 251  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-32676

Query Match  
Best Local Similarity 4.2%; Score 6; DB 4; Length 251;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 AKKVR 133  
Db 106 AKKVR 111

RESULT 74  
US-09-543-681A-6028  
; Sequence 6028, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-05  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 6028  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-6028

Query Match  
Best Local Similarity 4.2%; Score 6; DB 4; Length 252;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 AKKIA 114  
Db 60 AKKIA 65

RESULT 75  
US-09-252-991A-17964  
; Sequence 17964, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:

Tue Aug 17 05:54:49 2004

us-09-690-825-34.oligo.rat

Page 25

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 17964  
LENGTH: 260  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17964

Query Match 4.2%; Score 6; DB 4; Length 260;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 PPAWQP 12  
Db 45 PPAWQP 50

Search completed: August 11, 2004, 14:26:02  
Job time : 21 secs



Tue Aug 17 05:54:50 2004

us-09-690-825-34.oligo.rge

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 16, 2004, 00:14:49 ; Search time 3668 Seconds  
(without alignments)  
1668.648 Million cell updates/sec

Title: US-09-690-825-34

Sequence: 142  
1 NGAPILPMPKQPFILKDHRS.....EFFEPAKVRRAIRQLAAMD 142

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 2167151695 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6934737

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgm2/1/USPTO.spool/US09690825/runat\_11082004\_141014\_13791/app.query.fasta\_1.327  
-DB=GenEmbl -QFMT=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100  
-DOCALLIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=75 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US09690825.@CGN\_1\_1\_3731@runat\_11082004\_141014\_13791 -NCPV=6 -ICPV=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	142	100.0	426	6	AR184473
2	142	100.0	426	6	BD167854
3	142	100.0	426	6	BD185366
4	142	100.0	1619	6	AR099347
5	142	100.0	1619	6	BD273550
6	142	100.0	1619	6	AR181635
7	142	100.0	1619	6	AX775129
8	142	100.0	1619	6	AX779941
9	142	100.0	1629	9	AF077350
10	142	100.0	1643	9	BC034148
11	142	100.0	1653	9	BC008718
12	107	75.4	1011	9	HUMMRNASQ
13	107	75.4	1165	6	AR097642
14	107	75.4	1165	6	AR154345
15	106	74.6	1330	4	AB095108
16	93	65.5	740	6	AR270355
17	76	53.5	399	6	AX86452
18	76	53.5	399	6	BD026062
19	74	52.1	600	9	AB028869
20	73	51.4	1339	9	BC000784
21	68	47.9	463	6	AX778483
22	58	40.8	794	4	AP195781
23	41	28.9	924	10	AF077349
24	41	28.9	955	6	AR181548
25	41	28.9	955	10	AB013819
26	41	28.9	3352	10	BC004702
27	39	27.5	903	10	AF276775
28	39	27.5	14796	6	AR157507
29	39	27.5	14796	6	AR181541
30	39	27.5	14796	6	AX333491
31	39	27.5	14796	6	AX334154
32	39	27.5	14796	6	AX336483
33	39	27.5	14796	6	AX410775
34	39	27.5	14796	6	BD192864
35	39	27.5	14796	6	HSU75285
36	39	27.5	154840	2	AC032035
37	39	27.5	209751	9	AC087645
38	37	26.1	417	6	AR184472
39	37	26.1	154840	2	AC032035
40	37	26.1	229426	2	AC010532
41	34	23.9	779	10	CRUCIRA
42	34	23.9	9263	10	AF077351
43	34	23.9	10074	10	AB036741
44	34	23.9	10919	10	AF115517
45	34	23.9	197194	2	AL954690
46	34	23.9	203336	10	AL591433
47	20	14.1	143111	2	AC143833
48	19	13.4	138875	2	AC142802
49	19	13.4	186874	2	AC142803
50	15	10.6	229426	2	AC010532

AR184473 Sequence  
BD167854 Survivin-  
BD185366 Survivin-  
AR099347 Sequence  
BD273550 Antisense  
AR181635 Sequence  
AX775129 Sequence  
AX779941 Sequence  
AF077350 Homo sapi  
BC034148 Homo sapi  
BC008718 Homo sapi  
L26245 Human effec  
AR097642 Sequence  
AR154345 Sequence  
AB095108 Canis fam  
AR270355 Sequence  
AX86452 Sequence  
BD026062 Sequence  
AB028869 Homo sapi  
BC000784 Homo sapi  
AX778483 Sequence  
AP195781 Sus scrofa  
AF077349 Mus muscu  
AR181548 Sequence  
AB013819 Mus muscu  
BC004702 Mus muscu  
AF276775 Rattus no  
AR157507 Sequence  
AR181541 Sequence  
AX333491 Sequence  
AX334154 Sequence  
AX336483 Sequence  
AX410775 Sequence  
BD192864 Survivin,  
U75285 Homo sapien  
AC032035 Homo sapi  
AC087645 Homo sapi  
AR184472 Sequence  
AC032035 Homo sapi  
AC010532 Homo sapi  
W80243 Cricetus  
AF077351 Mus muscu  
AB036741 Mus muscu  
AF115517 Mus muscu  
AL954690 Mus muscu  
AL591433 Mouse DNA  
AC143833 Macaca mu  
AC142802 Macaca mu  
AC142803 Macaca mu  
AC010532 Homo sapi

51	13	9.2	772	5	AY174765
52	13	9.2	207347	2	AC141438 Mus muscu
53	12	8.5	483	5	AF442492 Xenopus 1
54	12	8.5	483	5	AF442492 Xenopus 1
55	12	8.5	503	5	AY115553 Xenopus 1
56	12	8.5	555	5	AF322051 Gallus ga
57	12	8.5	773	5	AF377323 Gallus ga
58	10	7.0	230800	10	AY100633 Xenopus 1
59	10	7.0	266342	2	ALB31742 Mouse DNA
60	9	6.3	30	6	AC109963 Rattus no
61	9	6.3	30	6	AX923830 Sequence
62	9	6.3	316	5	AX923830 Sequence
63	9	6.3	445	5	AF377322 Gallus ga
64	9	6.3	489	5	AF377322 Gallus ga
65	9	6.3	8883	5	AY057058 Danio rer
66	9	6.3	104435	9	AC107911 Homo sapi
67	9	6.3	136470	2	AC106292 Homo sapi
68	9	6.3	146515	2	AC145222 Medicago
69	9	6.3	180849	2	AC146981 Homo sapi
70	9	6.3	184592	2	BX649590 Danio rer
71	9	6.3	184717	9	AC009717 Homo sapi
72	9	6.3	188476	10	AC122238 Homo sapi
73	9	6.3	194837	10	AC122238 Homo sapi
74	9	6.3	199590	2	AL607066 Mouse DNA
75	9	6.3	202366	2	AL805939 Mus muscu
76	9	6.3	205117	2	AL805939 Mus muscu
77	9	6.3	220069	10	AL732470 Mouse DNA
78	9	6.3	226529	2	BX571968 Danio rer
79	9	6.3	228077	2	AC130552 Rattus no
80	9	6.3	240200	2	AC122075 Rattus no
81	9	6.3	261484	2	AC095477 Rattus no
82	9	6.3	323459	2	AC119555 Rattus no
83	8	5.6	26	6	AR181551 Sequence
84	8	5.6	293	6	AR421862 Sequence
85	8	5.6	293	6	BD117415 EST and e
86	8	5.6	326	6	AR421901 Sequence
87	8	5.6	326	6	BD117454 EST and e
88	8	5.6	337	6	AR421887 Sequence
89	8	5.6	337	6	BD117440 EST and e
90	8	5.6	342	6	AR421870 Sequence
91	8	5.6	342	6	BD117423 EST and e
92	8	5.6	346	6	AR421946 Sequence
93	8	5.6	346	6	BD117499 EST and e
94	8	5.6	349	6	AR421878 Sequence
95	8	5.6	349	6	BD117431 EST and e
96	8	5.6	351	6	AR421923 Sequence
97	8	5.6	351	6	BD117476 EST and e
98	8	5.6	357	6	AR421910 Sequence
99	8	5.6	357	6	BD117463 EST and e
100	8	5.6	361	6	AR421895 Sequence

## ALIGNMENTS

RESULT 1	AR184473	426 bp	DNA	linear	PAT 20-APR-2002
LOCUS	AR184473	Sequence 2 from patent US 6346389.			
DEFINITION	AR184473				
ACCESSION	AR184473.1	GI:20230438			
VERSION					
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 426)				
AUTHORS	Altieri, D.C.				
TITLE	Method for selectively modulating the interactions between survivin and tubulin				
JOURNAL	Patent: US 6346389-A 2 12-FEB-2002;				
FEATURES	Location/Qualifiers				
source	1..426				
	/organism="unknown"				
	/mol_type="unassigned DNA"				

ORIGIN					
Alignment Scores:					
Pred. No.:	2.1e-157	Length:	426		
Score:	142.00	Matches:	142		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-690-825-34 (1-142) x AR184473 (1-426)					
QY	1	MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer	20		
DB	1	ATGGGTGCCCGAGCGTTGCCCGCCCTGCGACGCCCTTCTCAAGACCAACCGATCTCT	60		
QY	21	ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu	40		
DB	61	ACATTCAAGAACTGGCCCTTCTTGGAGGCTGGCGCTGCACCCGAGCGGATGGCCGAG	120		
QY	41	AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys	60		
DB	121	GCTGGCTTATCCACCTGCCCGCCACTGAGAACGACCCAGCTGGCCCGAGTCTTCTTCC	180		
QY	61	PheLysGluLeuGluGlyTTPGluProAspAspAspProIleGluGluHisLysLysHis	80		
DB	181	TTCAAGAGCTGGAAGGCTGGAGCGCAATGACGACCCCATAGAGAACTAAAGACAT	240		
QY	81	SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGluGlu	100		
DB	241	TGCTCGGCTTGGCTTCTTCTTCTGTCAGAACGATTTGAGAAATTAACCTTGATGA	300		
QY	101	PheLeuLysLeuAspArgGluAlaGlyAlaLysAsnLysIleAlaLysGluThrAsnLys	120		
DB	301	TTTTTGAACCTGGACAGAAAGAGCCCAAGAACAAATTGCAAGGAAACCAACCAATTAG	360		
QY	121	LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaIa	140		
DB	361	AAGAAAGATTATAGAGAACTGGAGAAAGTGGCGCTGCTCATTCAGCAAGCTGCTGCC	420		
QY	141	MetAsp 142			
DB	421	ATGGAT 426			
RESULT 2	BD167854	426 bp	DNA	linear	PAT 17-JAN-2003
LOCUS	BD167854	Sequence 1 from patent US 6346389.			
DEFINITION	BD167854				
ACCESSION	BD167854.1	GI:27873666			
VERSION	WO 0233071-A/6.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 426)				
AUTHORS	Tanaka, H. and Kaieda, I.				
TITLE	Survivin-like polypeptide and its DNA				
JOURNAL	Patent: WO 0233071-A 6 25-APR-2002;				
COMMENT	TAKEDA CHEMICAL INDUSTRIES LTD, HIROSHI TANAKA, ISAO KAIEDA				
OS	Homo sapiens (human)				
PN	WO 0233071-A/6				
PD	25-APR-2002				
PF	16-OCT-2001 WO 2001JP009071				
PR	17-OCT-2000 JP 00P 316721, 20-DEC-2000 JP 00P 366809 PI				
HIROSHI TANAKA, ISAO KAIEDA					
PC	C12N15/09, C12N15/12, C07K14/47, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC				
10,	C12P21/02, C07K16/18, C12Q1/02, C12Q1/68, G01N33/15, G01N33/50, PC				
PC	A61K31/711, A61K39/395, A61K48/00, A61P35/00, A61P43/00 CC				
Survivin-like polypeptide and its DNA					

FEATURES

source

Location/Qualifiers

1..426

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

## ORIGIN

Alignment Scores:

Pred. No.: 2,1e-157 Length: 426

Score: 142.00 Matches: 142

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x BD167854 (1-426)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20

DB 1 ATGGGTGCCCGACGCTTCCCTGCTGGAGCCCTTTCTCAAGAGCACCGCATCTCT 60

QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40

DB 61 ACATTCAAGAACTGGCCCTTCTTGGAGGCGCTGCGCTCCACCCCGAGCGGATGCGGAG 120

QY 41 AlaGlyPheIleHisCysPProThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60

DB 121 GCTGGCTTCATCACTGACCCCACTGAGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 180

QY 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluGluHisLysLysHis 80

DB 181 TTCAGAGAGCTGGAAGCTGGAGGCGCAATGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 240

QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100

DB 241 TCGTCCGCTTGGCTTCTTCTGTCAGAGAGCAGTTGAAGATTACCTTTGCTGGA 300

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120

DB 301 TTTTGAACCTGACGAG 360

QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140

DB 361 AAGAAAGATTGGAGAACTGCGAAGAAAGTGGCCCTGCTCATGAGCACTGGCTGCC 420

QY 141 MetAsp 142

DB 421 ATGGAT 426

RESULT 3

BD185366 426 bp DNA linear PAT 17-JUN-2003

LOCUS Survivin-like polypeptide and its DNA.

DEFINITION BD185366

ACCESSION BD185366.1 GI:31877566

VERSION JP 2002355062-A/6.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 426)

1 Tanaka,H. and Kaleda,I.

AUTHORS Tanaka,H. and Kaleda,I.

TITLE Survivin-like polypeptide and its DNA

JOURNAL Patent: JP 2002355062-A 6 10-DEC-2002;

TAKEEDA CHEMICAL INDUSTRIES LTD

OS Homo sapiens (human)

PN JP 2002355062-A/6

PD 10-DEC-2002

PI 16-OCT-2001 JP 2001318533

PI HIROSHI TANAKA, ISAO KALEDA

PC C12N15/09,A61K31/7088,A61K38/55,A61K39/395,A61K39/395,A61K46/

PC 00,A61P35/00,

PC A61P43/00,A61P43/00,C07K14/82,C07K16/32,C12N1/15,C12N1/19, PC

C12N1/21,

PC C12N5/10,C12P21/02,C12Q1/02,C12Q1/68,G01N33/15,G01N33/50, PC

G01N33/53,

PC G01N33/53,G01N33/56,G01N33/574,C12N15/00,C12N5/00,A61K37/64

CC Survivin-like polypeptide and its DNA

PH key

FT source

FT source

Location/Qualifiers

1..426

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

## ORIGIN

Alignment Scores:

Pred. No.: 2,1e-157 Length: 426

Score: 142.00 Matches: 142

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x BD185366 (1-426)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20

DB 1 ATGGGTGCCCGACGCTTCCCTGCTGGAGCCCTTTCTCAAGAGCACCGCATCTCT 60

QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40

DB 61 ACATTCAAGAACTGGCCCTTCTTGGAGGCGCTGCGCTCCACCCCGAGCGGATGCGGAG 120

QY 41 AlaGlyPheIleHisCysPProThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60

DB 121 GCTGGCTTCATCACTGACCCCACTGAGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 180

QY 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluGluHisLysLysHis 80

DB 181 TTCAGAGAGCTGGAAGCTGGAGGCGCAATGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 240

QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100

DB 241 TCGTCCGCTTGGCTTCTTCTGTCAGAGAGCAGTTGAAGATTACCTTTGCTGGA 300

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120

DB 301 TTTTGAACCTGACGAG 360

QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140

DB 361 AAGAAAGATTGGAGAACTGCGAAGAAAGTGGCCCTGCTCATGAGCACTGGCTGCC 420

QY 141 MetAsp 142

DB 421 ATGGAT 426

RESULT 4

AR099347 1619 bp DNA linear PAT 14-FEB-2001

LOCUS Sequence 1 from patent US 6077709.

DEFINITION AR099347

ACCESSION AR099347.1 GI:12809113

VERSION AR099347.1

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1619)

1 Bennett,C.Frank., Ackermann,E.J., Swayze,E.E. and Cowsett,L.M.

AUTHORS Bennett,C.Frank., Ackermann,E.J., Swayze,E.E. and Cowsett,L.M.

TITLE Antisense modulation of Survivin expression

JOURNAL Patent: US 6077709-A 1-20-JUN-2000;  
Location/Qualifiers  
1..1619  
source /organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.: 7.81e-157 Length: 1619  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x AR099347 (1-1619)

QY 1 MetGlyAlaProThrLeuProProAlaATPGlnProPheLeuLysAspHisArgIleSer 20  
Db 50 ATGGGTCCTCCGACGCTGGCCCTGCTGACGCTTCTCTCAAGACACCGCATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 110 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGGCTGACCCCGAGGGGATGGCCGAG 169  
QY 41 AlaGlyPheLeuHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
Db 170 GCTGGCTTCATCCACGCTGCCCCCACTGAGAACGAGCAGACTGGCCAGTGTCTTCTGC 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisIleLysHis 80  
Db 230 TTCAAGAGCTGGAGAGCTGGAGAGCAGATGACGACCCCATAGAGAACATTAAGAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100  
Db 290 TCGTCCGCTTGGCGCTTCTTCTCTGTCAGAGCAGTTGAAGATTAACTTGTGTGA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
Db 350 TTTTGAACCTGGACGAGAAAGAGCCAGAACAAATTGCAAGAGAAACCAACATTAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140  
Db 410 AAGAAAGAAATTGAGAGAACTGCAAGAAAGTGGCCCGCATGACACACTGGCTGCC 469  
QY 141 MetAsp 142  
Db 470 ATGGAT 475

RESULT 5  
BD273550 1619 bp DNA linear PAT 17-JUL-2003  
LOCUS Antisense modulation of survivin expression.  
DEFINITION BD273550  
ACCESSION BD273550.1 GI:33083318  
VERSION JP 2002539073-A/1.  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1619)  
REFERENCE  
AUTHORS Bennett, F. C., Ackermann, E. J., Swazy, E. E. and Cowsett, L. M.  
TITLE Antisense modulation of survivin expression  
JOURNAL Patent: JP 2002539073-A 1 19-NOV-2002;  
OS Homo sapiens (human)  
IS PHARMACEUTICALS INC  
PN JP 2002539073-A/1  
PD 19-NOV-2002  
PF 23-SEP-1999 JP 2000572239  
PR 29-SEP-1998 US 09/163162, 05-APR-1999 US 09/286407 PI  
FRANK C BENNETT, ELIZABETH J ACKERMANN, ERIC E SWAZEY, LEX M PI  
COMSERT  
PC C07H21/04, A61K31/7088, A61K31/712, A61K48/00, A61J35/00 CC

Antisense modulation of survivin expression  
FH Key Location/Qualifiers  
FT CDS (50)..(478).  
source Location/Qualifiers  
1..1619  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 7.81e-157 Length: 1619  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x BD273550 (1-1619)

QY 1 MetGlyAlaProThrLeuProProAlaATPGlnProPheLeuLysAspHisArgIleSer 20  
Db 50 ATGGGTCCTCCGACGCTGGCCCTGCTGACGCTTCTCTCAAGACACCGCATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 110 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGGCTGACCCCGAGGGGATGGCCGAG 169  
QY 41 AlaGlyPheLeuHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
Db 170 GCTGGCTTCATCCACGCTGCCCCCACTGAGAACGAGCAGACTGGCCAGTGTCTTCTGC 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisIleLysHis 80  
Db 230 TTCAAGAGCTGGAGAGCTGGAGAGCAGATGACGACCCCATAGAGAACATTAAGAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100  
Db 290 TCGTCCGCTTGGCGCTTCTTCTCTGTCAGAGCAGTTGAAGATTAACTTGTGTGA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
Db 350 TTTTGAACCTGGACGAGAAAGAGCCAGAACAAATTGCAAGAGAAACCAACATTAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140  
Db 410 AAGAAAGAAATTGAGAGAACTGCAAGAAAGTGGCCCGCATGACACACTGGCTGCC 469  
QY 141 MetAsp 142  
Db 470 ATGGAT 475

RESULT 6  
AR181635 1619 bp DNA linear PAT 20-APR-2002  
LOCUS Antisense modulation of survivin expression.  
DEFINITION AR181635  
ACCESSION AR181635  
VERSION AR181635.1 GI:20223849  
KEYWORDS Unknown.  
SOURCE  
ORGANISM Unknown.  
Unclassified.  
REFERENCE  
AUTHORS Bennett, F. C., Ackermann, E. J., Swazy, E. E. and Cowsett, L. M.  
TITLE Antisense modulation of survivin expression  
JOURNAL Patent: US 635194-A 97 01-JAN-2002;  
OS Homo sapiens (human)  
IS PHARMACEUTICALS INC  
PN JP 2002539073-A/1  
PD 19-NOV-2002  
PF 23-SEP-1999 JP 2000572239  
PR 29-SEP-1998 US 09/163162, 05-APR-1999 US 09/286407 PI  
FRANK C BENNETT, ELIZABETH J ACKERMANN, ERIC E SWAZEY, LEX M PI  
COMSERT  
PC C07H21/04, A61K31/7088, A61K31/712, A61K48/00, A61J35/00 CC

ORIGIN

Alignment Scores:

US-09-690-825-34 (1-142) x BD273550 (1-1619)

QY 1 MetGlyAlaProThrLeuProProAlaATPGlnProPheLeuLysAspHisArgIleSer 20  
Db 50 ATGGGTCCTCCGACGCTGGCCCTGCTGACGCTTCTCTCAAGACACCGCATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 110 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGGCTGACCCCGAGGGGATGGCCGAG 169  
QY 41 AlaGlyPheLeuHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
Db 170 GCTGGCTTCATCCACGCTGCCCCCACTGAGAACGAGCAGACTGGCCAGTGTCTTCTGC 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisIleLysHis 80  
Db 230 TTCAAGAGCTGGAGAGCTGGAGAGCAGATGACGACCCCATAGAGAACATTAAGAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100  
Db 290 TCGTCCGCTTGGCGCTTCTTCTCTGTCAGAGCAGTTGAAGATTAACTTGTGTGA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
Db 350 TTTTGAACCTGGACGAGAAAGAGCCAGAACAAATTGCAAGAGAAACCAACATTAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140  
Db 410 AAGAAAGAAATTGAGAGAACTGCAAGAAAGTGGCCCGCATGACACACTGGCTGCC 469  
QY 141 MetAsp 142  
Db 470 ATGGAT 475



```

Pred. No.: 7.81e-157 Length: 1619
Score: 142.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-690-825-34 (1-142) x AR181635 (1-1619)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPhleuLysAspHisArgIleSer 20
Db 50 ATGGGTGCCCCGACGTTGCCCCCTGGCTGACGCTTTCTCAAGACCAACCGCATCTCT 109
QY 21 ThrPhleuAsnTrpProPhleuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 110 ACATTCAAGAACTGGCCCTTCTTGAGAGGCTGCCCTGCACCCCGAGCGGATGGCCGAG 169
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db 170 GCTGCTTCATCCACTGCCCCACAGACGACGACCTGGCCCTGCTTCTTCTGCTGC 229
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisLysLysHis 80
Db 230 TTCAGAGAGCTGGAAGGCTGGAGCCAGATGACGACCCCATAGAGAACATTAAGCAT 289
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100
Db 290 TCGTCCGGTTCGCTTCTTCTTCTGTCAGAGACGATTGAACCTTGGTGAA 349
QY 101 PheLysLysLeuAspArgGluThrAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
Db 350 TTTTGAACCTGACAGAGAAAGCCAGAAACAAATTGCAAGAGAAACCAACATTAAG 409
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140
Db 410 AAGAAAGAAATTGAGAGAACTGCCAGAGAAAGTGGCCGCTGCATCGAGCAGCTGCTGCC 469
QY 141 MetAsp 142
Db 470 ATGGAT 475

RESULT 7
AX775129 1619 bp DNA linear PAT 09-JUL-2003
LOCUS AX775129
DEFINITION Sequence 445 from Patent WO03038129.
ACCESSION AX775129
VERSION AX775129.1 GI:32486645
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1. Raponi, M.
TITLE Methods for assessing and treating leukemia
JOURNAL Patent: WO 03038129-A 445 08-MAY-2003;
Ortho-Clinical Diagnostics, Inc. (US)
FEATURES
source 1..1619
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 7.81e-157 Length: 1619
Pred. No.: 142.00 Matches: 142
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-09-690-825-34 (1-142) x AX775129 (1-1619)

```

```

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPhleuLysAspHisArgIleSer 20
Db 50 ATGGGTGCCCCGACGTTGCCCCCTGGCTGACGCTTTCTCAAGACCAACCGCATCTCT 109
QY 21 ThrPhleuAsnTrpProPhleuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 110 ACATTCAAGAACTGGCCCTTCTTGAGAGGCTGCCCTGCACCCCGAGCGGATGGCCGAG 169
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db 170 GCTGCTTCATCCACTGCCCCACAGACGACGACCTGGCCCTGCTTCTTCTGCTGC 229
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisLysLysHis 80
Db 230 TTCAGAGAGCTGGAAGGCTGGAGCCAGATGACGACCCCATAGAGAACATTAAGCAT 289
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100
Db 290 TCGTCCGGTTCGCTTCTTCTTCTGTCAGAGACGATTGAAGATTAACCTTGGTGAA 349
QY 101 PheLysLysLeuAspArgGluThrAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
Db 350 TTTTGAACCTGACAGAGAAAGCCAGAAACAAATTGCAAGAGAAACCAACATTAAG 409
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140
Db 410 AAGAAAGAAATTGAGAGAACTGCCAGAGAAAGTGGCCGCTGCATCGAGCAGCTGCTGCC 469
QY 141 MetAsp 142
Db 470 ATGGAT 475

RESULT 8
AX779941 1619 bp DNA linear PAT 14-JUL-2003
LOCUS AX779941
DEFINITION Sequence 2098 from Patent WO03039443.
ACCESSION AX779941
VERSION AX779941.1 GI:32696935
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1. Haeflrich, T., Schoch, C., Kern, W., Kohlmann, A., Schmittger, S.,
Dugas, M., Bils, R., Broze, B. and Wergenthaler, S.
TITLE Novel genetic markers for Leukemias
JOURNAL Patent: WO 03039443-A 2098 15-MAY-2003;
Deutsches Krebsforschungszentrum (DE);
Ludwig-Maximilians-Universitaet Muenchen (DE);
Hafelrich, T., Schoch, Claudia (DE); Kern, Wolfgang (DE)
FEATURES
source 1..1619
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 7.81e-157 Length: 1619
Pred. No.: 142.00 Matches: 142
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-09-690-825-34 (1-142) x AX779941 (1-1619)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPhleuLysAspHisArgIleSer 20
Db 50 ATGGGTGCCCCGACGTTGCCCCCTGGCTGACGCTTTCTCAAGACCAACCGCATCTCT 109

```



JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 MEDLINE 22388257  
 PUBMED 12477932  
 REFERENCE 2 (bases 1 to 1643)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-JUL-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing By: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcg@paxil.stanford.edu](mailto:mcg@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IPAL Plate: 41 Row: h Column: 21.

## FEATURES

## SOURCE

## Location/Qualifiers

1..1643  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:32768 IMAGE:4656567"  
 /tissue\_type="Lung, mucopidermoid carcinoma"  
 /clone\_id="NTH\_MGC\_59"  
 /lab\_host="DH10B"  
 /note="Vector: pDNR-LIB"  
 1..1643  
 /gene="BIRC5"  
 /note="synonyms: API4, SURVIVIN, EPR-1"  
 /db\_xref="LocusID:333"  
 /db\_xref="WIM:603352"  
 45..473  
 /codon\_start=1  
 /product="baculoviral IAP repeat-containing protein 5"  
 /protein\_id="AAH34148.1"  
 /db\_xref="GI:21707887"  
 /db\_xref="LocusID:333"  
 /translation="MGAPLTPANQPLKDHRISTPKNMFLEGCACCTPERMAEAGFI  
 HCFTEHPDLACFCPCFKELEGWEPDDPIEHKKHSGCAFLSKVQFELITGEFL  
 KUDREKANKAKETNNKKKEFEETAKKVRRAIEQLAAMD"  
 84..305  
 /note="3'IR; Region: Baculoviral inhibition of apoptosis  
 protein repeat"  
 /db\_xref="CDD:smat00238"

## gene

## CDS

## misc\_feature

84..305  
 /note="3'IR; Region: Baculoviral inhibition of apoptosis  
 protein repeat"  
 /db\_xref="CDD:smat00238"

## ORIGIN

## Alignment Scores:

Pred. No.: 7.92e-157 Length: 1643  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x BC034148 (1-1643)

QY 1 MetGtAlAPrThreupProAlaATrGInProPhelLeuLyAspHisArgLieser 20  
 DB 45 ATGGGAGCCCGACGCTGGCCCTGGCTGGACGCTTTCTCAAGACACCGCATCTCT 104  
 QY 21 ThrPhelYasnTrpProPhelLeuGluGlyCYsAlaCYsThProGluArgMetLlaGlu 40  
 DB 105 ACATTCAGAAAGTGGCCCTTCTTGGAGGGCTGGCTGGACCGGAGGATGGCCGAG 164

QY 41 AlaGlyPheIleHisCYsProThrcGluAsnGluProAspLeuAlaGlnCYsPhePheCYs 60  
 DB 165 GCTGGCTTCATTCATCTGCCCTCCCTGAGAACGAGCCAGACTGGCCCATGTTCTTCGCG 224  
 QY 61 PheLYsGluLeuGluGlyTYrGluProAspAspProIleGluGlnHisLYsHis 80  
 DB 225 TTCAGAGAGCTGAGAGCTGGAGCCAGATGCGACCCCATAGAGAAATATAAAAGCAT 284  
 QY 81 SerSerGlyCYsAlaPheLeuSerValLYsGlnPheGluGluLeuThrcGluGlyGlu 100  
 DB 285 TCGTCGGTGGCTGCTTCTTCCTTCTCTCAAGAGCGATTGAGAAATTAACCTTGCTGTA 344  
 QY 101 PheLeuLYsLeuAspArgGluArgAlaLYsAsnLYsIleAlaLYsGluThrAsnAsnLYs 120  
 DB 345 TTTTGGAAACTGCAAGAGAAAGACCAAGAAATTCAGAAAGAAACCAATATAG 404  
 QY 121 LYsLYsGluPheGluGluGluThrAlaLYsLYsValArgArgAlaIleGluGlnLeuAla 140  
 DB 405 AAGAAAGAAATTTGAGAAACTGCGAAGAAAGTGGCCGTCATGACGACTGGCTGCG 464  
 QY 141 MetAsp 142  
 DB 465 ATGGAT 470

## RESULT 11

## BC008718

## LOCUS

## DEFINITION

Homo sapiens baculoviral IAP repeat-containing 5 (survivin), mRNA  
 (CDNA clone MGC:8592 IMAGE:296114), complete cds.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

1 (bases 1 to 1653)  
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Sapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Schenck, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
 Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullany, S.J., Bosak, S.A., McSwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, Y.K., S.W.,  
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shchepochko, Y.,  
 Sanchez, A., Whiting, M., Madan, A., Green, E.D.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, D.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,  
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 MEDLINE 22388257  
 PUBMED 12477932  
 REFERENCE 2 (bases 1 to 1653)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-MAY-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT On Nov 6, 2003 this sequence version replaced gi:14250533.  
 Contact: MGC help desk  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland:  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: nisc\_mgcmgr1.nih.gov  
 Ahter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Gouite, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Larc, P., Legaspi, R.,  
 Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, D.W.,  
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>  
 Series: IRAL Plate: 1 Row: 1 Column: 7  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 5453604.

## FEATURES

## source

Location/Qualifiers  
 1..1653  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:8592 IMAGE:2961114"  
 /tissue\_type="Muscle, rhabdomyosarcoma"  
 /clone\_id="NH\_MGC\_17"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 1..1653  
 /gene="BIRC5"  
 /note="synonyms: API4, SURVIVIN, EPR-1"  
 /db\_xref="LocusID:332"  
 /db\_xref="MIM:603352"  
 34..462  
 /codon\_start=1  
 /product="baculoviral IAP repeat-containing protein 5"  
 /protein\_id="AAH08718.1"  
 /db\_xref="GI:14250534"  
 /db\_xref="LocusID:332"  
 /translation="MGAPFLPPAPQPFILKDRISTFKWPFLEGCATPERMAAGFI  
 HCPTENEPDLACFCFKELBCEPDDEPDEBKHGSSGCAFLVKQFELTIGBFL  
 KIDRRARAKIKAKETNNKKKEEETAKKVRRAIEQLAMD"

## gene

## CDS

misc\_feature  
 73..294  
 /note="BIR: Region: Baculoviral inhibition of apoptosis  
 protein repeat"  
 /db\_xref="CDD:smart00238"

## ORIGIN

Alignment Scores:  
 Pred. No.: 7.97e-157 Length: 1653  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x BC008718 (1-1653)

QY 1 MetGlyAlaProThleuProProAlaTrpGlnProPhelenuyAspHisArgIleSer 20  
 Db 34 ATGGGAGCCCGACGCTGCGCCCTGCGAGCCCTTTCTCAAGAGACACCGCATCTCT 93  
 QY 21 ThrPhelyAsnTrpProPhelenuyGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 94 ACATTCAAGAACTGGCCCTCTTGTGAGGGCTGCGCTGCACCCGAGCGGATGCGCGAG 153  
 QY 41 AlaGlyPheIleHisCysProThrGluGlnGluProAbelLysAlaGlnCysPhePheCys 60  
 Db 154 GCTGCTTCATCCTACCTGCGCCCACTGAGAACGAGCAGCATGCGCCAGTGTCTCTGCG 213

QY 61 PheLysGluLeuGluGlyTyrGluProAspAspProIleGluGluHisLysHis 80  
 Db 214 TTCAAGAGCTGGAGGCTGCGAGCAGATGACGCCCATAGAGAACTAAAGGCAAT 273  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysGluGlnPheGluGluThrLeuGlyGlu 100  
 Db 274 TCGCCGCTGGCCCTTCTCTTCTGTCGAAGACAGCTTGAAATAATACCTTGGTGA 333  
 QY 101 PheLeuLysLeuAspArgGluAArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 Db 334 TTTTGAATGAGCAGAGAGAAAGCCAGAAACAAATTGCAGAAAGAAACCAATAG 393  
 QY 121 LysLysGluLeuGluGlyTyrAlaLysValLysValArgArgAlaIleGluGluLeuAla 140  
 Db 394 AAGAAAGAAATTGAGGAACTGGAAGAAAGTGGCGGCGCATCGACGAGCTGGCGC 453

## RESULT 12

## HUMMRNASEQ/c

LOCUS HUMMRNASEQ 1011 bp mRNA linear PRI 13-JUL-1994  
 DEFINITION Human effector cell protease receptor-1 (EPR-1) mRNA, partial cds.  
 ACCESSION L26245  
 VERSION L26245.1 GI:456089

## KEYWORDS

## SOURCE

## ORGANISM

Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## FEATURES

## source

## Location/Qualifiers

1..1011  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_id="lambda gt11 MLT, gt10 HEK, gt11 HUVEC"  
 1..1011  
 /gene="EPR-1"  
 1..>1011  
 /gene="EPR-1"  
 /codon\_start=1  
 /product="effector cell protease receptor 1"  
 /protein\_id="AA19687.1"  
 /db\_xref="GI:456090"  
 /translation="MTSRGFORSCNNPSSSVTGRANOIHLTPDFISRELLPEPKA  
 GTWACVSPGCEGRDCEGMADRHACSSPACTCOVHODCSLNMMSSRTHGCRIC  
 HANKAVSSSKRDTRAPFLPHFSCKFGONSKRBPXKREBSHCYRAAPAKQAEAP  
 LVPLSRONKSTIVETSNLKYLLISFPRTLNGPQEGWHDGINSQAAVTLGPPSSSRPQ  
 STBASQMARRTFPAVSSNSPFLVLSFALIFLALISLSPKNSPRVNSNCTFLERKA  
 QPEDEGFLCSWSSGSSGSPSSSLIKQKKHMAKSGSPSVQMMKPSAIAISGVQRSPPR  
 AAS"

## CDS

ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.69e-116 Length: 1011  
 Score: 107.00 Matches: 107  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 75.35% Indels: 0  
 DB: 9 Gaps: 0

## ORIGIN

US-09-690-825-34 (1-142) x HUMMRNASEQ (1-1011)

QY 33 CysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrGluAsnGluPro 52

Db 983 TGCACCCCGAGGAGATGCCGAGGCTGCTTCATCCACTGCCCACTGAGAACGACCA 924

Qy 53 AspleuA1aGlnCysPhePheCysPheLysGluLeuGluGlyTTPGluProAspAspAP 72

Db 923 GACTGGCCCGAGGTTTCTTCTCTCCAGAGCTGAGAGCTGGAGCGCAGATGACGAC 864

Qy 73 ProLleGluGluHisLysLysHisSerSerGlyCysAlaPheLeuSerValLysLysGln 92

Db 863 CCCTAAGAGAACATTAAGACATTCCTCCGCTTCCCTTCTCTCTCAAGAAAGCAG 804

Qy 93 PheGluGluLeuThrLeuGlyGlyPheLeuLysLeuAspArgGluArgAlaLysAsnLys 112

Db 803 TTTCAGAAATTAACCTTGTGTGATTTTGAACCTGACAGAGAAAGACCAAGAAACAA 744

Qy 113 IleAlaLysGluThrAsnAsnLysLysLysGluPheGluGluThrAlaLysValArg 132

Db 743 ATTGCAAGAGAACCAATTAAGAACAAAGATTGAGAAACTCGAAGAAAGTGCCG 684

Qy 133 ArgAlaLleGluGlnLeuAla 139

Db 683 CGTCCATCGAGCGCTGGCC 663

RESULT 13

AR097642/c 1165 bp DNA linear PAT 14-FEB-2001

LOCUS AR097642

DEFINITION Sequence 1 from patent US 6072028.

ACCESSION AR097642

VERSION AR097642.1 GI:12806372

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1165)

AUTORS Altieri,D.C.

TITLE EPR-1 proteins, polypeptides, and nucleic acid molecules encoding same

JOURNAL Patent: US 6072028-A 1 06-JUN-2000;

FEATURES

source

1. 1165

/organism="unknown"

/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.: 8,85e-116 Length: 1165

Score: 107.00 Matches: 138

Percent Similarity: 99.28% Conservative: 0

Best Local Similarity: 99.28% Mismatches: 1

Query Match: 75.35% Indels: 1

Gaps: 0

US-09-690-825-34 (1-142) x AR097642 (1-1165)

Qy 1 MetGlyAlaProThrLeuProProAlaTTPGlnProPheLeuLysAspHisArgLleSer 20

Db 1114 ATGGTGCCCCGAGCTTGCCCTGCTGCGAGCCCTTCTCAAGAGACACCGCATCTCT 1055

Qy 21 ThrPheLysAsnTTPProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40

Db 1054 ACATTCAGAACTGCCCCCTTCTTGAGAGGCTGCGC-TGCACCCCGAGCGAGTGCAG 996

Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60

Db 995 GGTGGCTTCATCATGCCCCCCTCAGAACAGCAGCACTGGCCCACTGTTCTTCTGCG 936

Qy 61 PheLysGluLeuGluGlyTTPGluProAspAspAspProLleGluGluHisLysHis 80

Db 935 TTCAGAGAGCTGGAAGGCTGGAGCCAGATGACGCCCATAGAGAACTAAGAAAGCAT 876

Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100

Db 875 TCGTCCGTTGCGCTTCTTCTCTGTCAGAGAGCAGTTTGAAGAAATTAACCTTGTTGA 816

Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120

Db 815 TTTTGAAGACTGACAGAGAAAGAGCCCAAGAAACAAATTTGCAAGAGAAACCAACATTAAG 756

Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaLleGluGlnLeuAla 139

Db 755 AAGAAAGAAATTTAGAGAAACTGCGAGAAAGTGCCTGCTGCTGACGACAGCTGGCC 699

RESULT 14

AR154245/c 1165 bp DNA linear PAT 08-AUG-2001

LOCUS AR154245

DEFINITION Sequence 1 from patent US 6238875.

ACCESSION AR154245

VERSION AR154245.1 GI:15122298

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1165)

AUTORS Altieri,D.C.

TITLE Diagnostic methods useful in the characterization of lymphoproliferative disease characterized by increased EPR-1

JOURNAL Patent: US 6238875-A 1 29-MAY-2001;

FEATURES

source

1. 1165

/organism="unknown"

/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.: 8,85e-116 Length: 1165

Score: 107.00 Matches: 138

Percent Similarity: 99.28% Conservative: 0

Best Local Similarity: 99.28% Mismatches: 1

Query Match: 75.35% Indels: 1

Gaps: 0

US-09-690-825-34 (1-142) x AR154245 (1-1165)

Qy 1 MetGlyAlaProThrLeuProProAlaTTPGlnProPheLeuLysAspHisArgLleSer 20

Db 1114 ATGGTGCCCCGAGCTTGCCCTGCTGCGAGCCCTTCTCAAGAGACACCGCATCTCT 1055

Qy 21 ThrPheLysAsnTTPProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40

Db 1054 ACATTCAGAACTGCCCCCTTCTTGAGAGGCTGCGC-TGCACCCCGAGCGAGTGCAG 996

Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60

Db 995 GGTGGCTTCATCATGCCCCCCTCAGAACAGCAGCACTGGCCCACTGTTCTTCTGCG 936

Qy 61 PheLysGluLeuGluGlyTTPGluProAspAspAspProLleGluGluHisLysHis 80

Db 935 TTCAGAGAGCTGGAAGGCTGGAGCCAGATGACGCCCATAGAGAACTAAGAAAGCAT 876

Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100

Db 875 TCGTCCGTTGCGCTTCTTCTCTGTCAGAGAGCAGTTTGAAGAAATTAACCTTGTTGA 816

Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120

Db 815 TTTTGAAGACTGACAGAGAAAGAGCCCAAGAAACAAATTTGCAAGAGAAACCAACATTAAG 756

Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaLleGluGlnLeuAla 139

Db 755 AAGAAAGAAATTTAGAGAAACTGCGAGAAAGTGCCTGCTGCTGACGACAGCTGGCC 699

RESULT 15

AB095108 1630 bp mRNA linear MAM 27-NOV-2002

LOCUS AB095108

DEFINITION Canis familiaris mRNA for survivin, complete cds.

ACCESSION AB095108

VERSION AB095108.1 GI:24636590



REFERENCE 1  
AUTHORS Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.  
TITLE Expressed sequence tags and encoded human proteins  
JOURNAL Patent: EP 1033401-A 2315 06-SEP-2000;  
Genet (FR)  
FEATURES  
source Location/Qualifiers  
1..399  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
52..399  
/note="unassigned protein product"  
/codon\_start=1  
/protein\_id="CA00590.1"  
/db\_xref="GI:40043628"  
/translation="MGAPLTPAMQPLKDHRISTKMPFLGCACTPERNAAGPI  
HPTENPDIAOCFPCFKELEWEPDDDDPIEEKKHSSCAFSLVKQFEELTIGFEL  
KLDREKRNKRIAKE"  
ORIGIN  
Alignment Scores:  
Pred. No.: 9,46e-80 Length: 399  
Score: 76.00 Matches: 115  
Percent Similarity: 98.29% Conservative: 0  
Best Local Similarity: 98.29% Mismatches: 1  
Query Match: 53.52% Indels: 2  
DB: Gaps: 0  
US-09-690-825-34 (1-142) x AX886452 (1-399)  
QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 52 ATGGGTCCTCCGACGCTTCCCTCCCTGCGACGCTTCTTCAAGACACCGCATCTCT 111  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 112 ACATTCAAGAACTGGCCCTCTTGGAGGGCTGCGCCCTGCAACCCGAGCGGATGGCCGAG 171  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
Db 172 GCTGCTTCATCCATCCGACCTGAGAACGACGACCTTGGCCCACTGTTCTTCTGCG 231  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHis-LysLysHis 80  
Db 232 TTCAAGAGAGCTGGAGGCTGGAGCCAGATGACGACCCCATAGAGAGASA-TAAAAAGCA 290  
QY 80 sSerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGluGly 100  
Db 291 TTCGTCGGGTGGCGCTTCTTCTTGTSAAGAGAGGTTTGAAGATTAACTGGTGA 350  
QY 100 uPheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGlu 116  
Db 351 ATTTTGAACCTGACAGAGAAAGAGCCAGAAACAATTGCAAGAGAA 399  
RESULT 18  
LOCUS BD026062 399 bp DNA linear PAT 27-AUG-2002  
DEFINITION Sequence tag and encoded human protein.  
ACCESSION BD026062  
VERSION BD026062.1 GI:22567285  
KEYWORDS UP 2001269182-A/2308.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.  
TITLE Sequence tag and encoded human protein.  
JOURNAL Patent: JP 2001269182-A 2308 02-OCT-2001;  
GENSET  
COMMENT OS Homo sapiens (human)  
PN JP 2001269182-A/2308  
PD 02-OCT-2001

PF 24-FEB-2000 JP 2000118773  
PR 26-FEB-1999 US 60/122487  
PI JEAN BAPTIST DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES  
PJ JORDAN  
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC  
C12N5/10,  
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC  
G06F15/40  
FEATURES  
source Location/Qualifiers  
FT CDS  
1..399  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 9,46e-80 Length: 399  
Score: 76.00 Matches: 115  
Percent Similarity: 98.29% Conservative: 0  
Best Local Similarity: 98.29% Mismatches: 1  
Query Match: 53.52% Indels: 2  
DB: Gaps: 0  
US-09-690-825-34 (1-142) x BD026062 (1-399)  
QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 52 ATGGGTCCTCCGACGCTTCCCTCCCTGCGACGCTTCTTCAAGACACCGCATCTCT 111  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 112 ACATTCAAGAACTGGCCCTCTTGGAGGGCTGCGCCCTGCAACCCGAGCGGATGGCCGAG 171  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
Db 172 GCTGCTTCATCCATCCGACCTGAGAACGACGACGACCTTGGCCCACTGTTCTTCTGCG 231  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHis-LysLysHis 80  
Db 232 TTCAAGAGAGCTGGAGGCTGGAGCCAGATGACGACCCCATAGAGAGASA-TAAAAAGCA 290  
QY 80 sSerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGluGly 100  
Db 291 TTCGTCGGGTGGCGCTTCTTCTTGTSAAGAGAGGTTTGAAGATTAACTGGTGA 350  
QY 100 uPheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGlu 116  
Db 351 ATTTTGAACCTGACAGAGAAAGAGCCAGAAACAATTGCAAGAGAA 399  
RESULT 19  
LOCUS AB028869 600 bp mRNA linear PRI 04-APR-2000  
DEFINITION Homo sapiens mRNA for survivin-beta, complete cds.  
ACCESSION AB028869  
VERSION AB028869.1 GI:7416052  
KEYWORDS survivin-beta.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Kageyama, H., Islam, A., Takayasu, H. and Nakagawara, A.  
TITLE An isoform of Survivin (Survivin-beta) which has 23 Amino Acids  
JOURNAL Insertion into the BIR Domain  
Unpublished  
2 (bases 1 to 600)  
Kageyama, H., Islam, A. and Nakagawara, A.  
Direct Submission  
Submitted (11-JUN-1999) Akira Nakagawara, Chiba Cancer Center  
Research Institute, Division of Biochemistry, 666-2 Nipona,

Chuch-Ku, Chiba, China 260-8717, Japan  
E-mail: akiranak@chiba-cs.pref.chiba.jp,  
Tel: 81-43-264-5431 (ex. 5201), Fax: 81-43-265-4459)  
Location/Qualifiers

## FEATURES

1..600  
/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/cell\_line="CHP134"

/cell\_type="neuroblastoma"

27..524  
/codon\_start=1

/product="survivin-beta"

/protein\_id="BAA93676.1"

/db\_xref="GI:7416053"

/translation="MSAPLPPAPKPPFLKDRISTFQKWPFLBGCATPPEMAAGFI  
HCPTEHPDLAQCFCEKLEGEPPDDPIGATVAACNTSLGGRGRTTREHKK  
HSSGAFSLSVKKQFEBELTLGEPLKIDREARKKIKETNNKKPEFEYAKIVRAIEQ  
LAAMD"

## ORIGIN

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
74.00	100.00%	600	74
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	52.11%	Indels:	0
DB:	9	Gaps:	0

US-09-690-825-34 (1-142) x AB028869 (1-600)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20

Db 27 ATGGGTCGCCCGACGCTGGCCCTGGCTGGACGCCCTTCTCAAGACACCGCATCTCT 86

QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaIle 40

Db 87 ACATTAAAGAACTGGCCCTCTTGTGAGGGCTGCGCTCGACCCCGAGCGGATGGCCAG 146

QY 41 AlaGlyPheIleHisCysProThrGluGlnGluProAspPheLysAlaGlnCysPhePheCys 60

Db 147 GCTGGCTTATCATCATCGCCCACTGAGAACGAGCCAGACTTGCCCACTGTTCTTCTGC 206

QY 61 PheLysGluLeuGluGluGlyTrpGluProAspAspProIle 74

Db 207 TTCAGAGAGCTGGAAGGCTGGAGGACGATGACGACCCCAT 248

RESULT 20

BC000784 1539 bp mRNA linear PRI 16-SEP-2003

LOCUS BC000784

DEFINITION Homo sapiens baculoviral IAP repeat-containing 5 (survivin), mRNA

ACCESSION BC000784

VERSION BC000784.2 GI:34783885

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1539)

Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G.,

Klausner, R. D., Collins, F. S., Wagner, J., Shennan, C. M., Schuler, G. D.,

Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K.,

Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J. F., Heien, F.,

Diatchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L.,

Stapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L.,

Scheetz, T. E., Brownstein, M. J., Uedl, T. B., Toshiyuki, S.,

Carinini, P., Prange, C., Rana, S. S., Loguigliano, N. A., Peters, G. J.,

Abrahamson, R. D., Mullaly, S. J., Bosak, S. A., McEwan, P. J.,

McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S.,

Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. M.,

Villalobon, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A.,

Faney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## AUTHORS

## JOURNAL

## REMARK

## COMMENT

## FEATURES

## source

## 1..1539

## /organism="Homo sapiens"

## /mol\_type="mRNA"

## /db\_xref="taxon:9606"

## /clone="IMAGE:3506845"

## /tissue\_type="kidney, renal cell adenocarcinoma"

## /clone\_id="NIH-MGC\_14"

## /lab\_host="DH10B-R"

## /note="Vector: pOTB7"

## ORIGIN

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
1.21e-75	73.00	1539	73
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	51.41%	Indels:	0
DB:	9	Gaps:	0

US-09-690-825-34 (1-142) x BC000784 (1-1539)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20

Db 46 ATGGGTCGCCCGACGCTGGCCCTGGCTGGACGCCCTTCTCAAGACACCGCATCTCT 105

QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaIle 40

Db 106 ACATTAAAGAACTGGCCCTCTTGTGAGGGCTGCGCTCGACCCCGAGCGGATGGCCAG 165

QY 41 AlaGlyPheIleHisCysProThrGluGlnGluProAspPheLysAlaGlnCysPhePheCys 60

Db 166 GCTGGCTTATCATCATCGCCCACTGAGAACGAGCCAGACTTGCCCACTGTTCTTCTGC 225

QY 61 PheLysGluLeuGluGluGlyTrpGluProAspAspProIle 73

Db 226 TTCAGAGAGCTGGAAGGCTGGAGGACGATGACGACCC 264

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAL Plate: 7 Row: c Column: 10  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4502144.

Location/Qualifiers

1..1539

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3506845"

/tissue\_type="kidney, renal cell adenocarcinoma"

/clone\_id="NIH-MGC\_14"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

Sanchez, A., Whiting, N., Madan, A., Young, A. C., Shevchenko, Y.,  
Boutfield, G. G., Blakesley, R. W., Touchman, D. W., Green, E. D.,  
Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M.,  
Butterfield, Y. S., Krzyzanski, M. I., Skalska, U., Small, D. B.,  
Schnerch, A., Schein, J. E., Jones, S. J., and Marra, M. A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 1539)

Strausberg, R.

Direct Submission

Submitted (15-NOV-2000) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Sep 16, 2003 this sequence version replaced gi:12653976.

Contact: MGC help desk

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: DCTD/DRP

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

Contact: [amadnan@systemsbiology.org](mailto:amadnan@systemsbiology.org)

Anup Madan, Jessica Fahney, Erin Helton, Mark Kettelman, Anuradha  
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting



JOURNAL	Submitted (18-Oct-1999) Animal Science and Animal Health, The Royal Veterinary and Agricultural University, Groenegeardsvej 3, Frederiksberg C, 1870, Denmark		
FEATURES	Location/Qualifiers		
source	1..794		
	/organism="Sus scrofa"		
	/mol_type="mRNA"		
	/db_xref="taxon:9823"		
	/chromosome="12"		
	/map="12q11-q15; ETR-1 encoded by complement strand"		
	6..434		
	/note="contains 1 BIR and 1 COOH-terminus coiled domain"		
	/codon_start=1		
	/product="apoptosis inhibitor survivin"		
	/protein_id="AA037540.1"		
	/db_xref="GI:10443628"		
	/translation="MSAPSPAPWOLYTKDHRISTPKNMPLEGCACTPERMAAGPFIHCPENEDLACCFCKPEKLEGEWEPDDDIIEHKHSGGCAITVKQFESTLTSEFLKIDPERAKNKIAETNNKQKEFEETAKVRCALIEQLAASE"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,866-58	Length:	794
Score:	58.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	40.85%	Indels:	0
DB:	4	Gaps:	0
US-09-690-825-34 (1-142) x AF195761 (1-794)			
QY	41	AlAGlyPheIleHsCYsPProThcLIAsnGluPProAspleuAlaGInCYsPhePecYs	60
Db	126	GCGGGTTCATCCACTGCTCCACAGAAAGAGCCGATTTGGCTCAAGTTCTTCTTCG	185
QY	61	PheLYsGluLeuGluGlyTTRGluPProAspAspProIleGluGluHisLYsLYsHis	80
Db	186	TTCAAGAGACTGGAGGCTGGAGACCAGACGACGCCCTATGAGAAATATAAAGCT	245
QY	81	SeTSerGlyCYsAlaPheLeuSeRValLYsLYsGlnPheGInGluLeuThreU	98
Db	246	TCTCCGGTGGTTCCTTCTTGTGTAAGAGAGAGATTGAAAGATTAAACCTC	299
RESULT 23			
AF077349	924 bp mRNA linear ROD 14-DEC-2000		
LOCUS	AF077349 Mus musculus inhibitor of apoptosis homolog mRNA, complete cds.		
DEFINITION	AF077349		
ACCESSION	AF077349.1 GI:4959076		
VERSION			
KEYWORDS	Mus musculus (house mouse)		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognath; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 924)		
TITLE	Uren,A.G., Wong,L., Pakusch,M., Fowler,K.J., Burrows,F.J., Vaux,D.L. and Choo,K.H.		
JOURNAL	Survivin and the inner centromere protein INCENP show similar cell-cycle localization and gene knockout phenotype		
MEDLINE	Curr. Biol. 10 (21), 1319-1328 (2000)		
PUBMED	20538917		
REFERENCE	1 (bases 1 to 924)		
AUTHORS	Uren,A.G. and Vaux,D.L.		
TITLE	Direct Submision		
JOURNAL	Submitted (10-SEP-1997) Molecular Cancer Division, The Walter and Eliza Hall Institute, Royal Parade, Parkville, Vic 3052, Australia		
FEATURES	Location/Qualifiers		
source	1..924		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
	/clone="IMAGE Consortium clone ID 422303"		

CDS  
/dev\_stage="embryo; E13.5 - E14.5"  
16..438  
/note="IAP; contains single baculovirus IAP repeat (BIR)  
motif"  
/codon\_start=1  
/product="inhibitor of apoptosis homolog"  
/protein\_id="AAD34225.1"  
/db\_xref="GI:4959077"  
/translation="MGAPALPOLIOLYLNRYIAFKWMPLEDCATPERMAAGFI  
HCPTEHPDLAOCFFCFKELBEMEPDNDIEBHRKHSRSCAFITVKOMEELTVSEFL  
KIDROPARKKIATKNTNNKKEFEETAKTTRQSI EQLAA"

ORIGIN  
Alignment Scores:  
Pred. No.: 3,39e-38 Length: 924  
Score: 41.00 Matches: 41  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 28.87% Indels: 0  
DB: 10 Gaps: 0

US-09-690-825-34 (1-142) x AF077349 (1-924)

QY 31 CysAlaCysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrgluAsn 50  
DB 106 TGGCGCTGCACCCGAGAGGAATGGCGAGGCTGCTTCATCCTCACTGCCCTACCGAGAAC 165  
QY 51 GluProAspLeuAlaGlnCysPhePheCysPheIleGluGluGlyTyrGluProasp 70  
DB 166 GAGCCGATTGGCCCAAGTGTTCCTTTCCTTAAGAAATTGGAAGCTGGGAACCCGAT 225  
QY 71 Asp 71  
DB 226 GAC 228

RESULT 24  
AR181548  
LOCUS AR181548 955 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 10 from patent US 6335194.  
ACCESSION AR181548  
VERSION AR181548.1 GI:20223762  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 955)  
AUTHORS Bennett,C.,Frank., Ackermann,E.J., Swayze,E.E. and Cowsett,L.M.  
TITLE Antisense modulation of survivin expression  
JOURNAL Patent: US 6335194-A 10 01-JAN-2002;  
FEATURES  
source 1..955  
/organism="Unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.5e-38 Length: 955  
Score: 41.00 Matches: 41  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 28.87% Indels: 0  
DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x AR181548 (1-955)

QY 31 CysAlaCysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrgluAsn 50  
DB 199 TGGCGCTGCACCCGAGAGGAATGGCGAGGCTGCTTCATCCTCACTGCCCTACCGAGAAC 258  
QY 51 GluProAspLeuAlaGlnCysPhePheCysPheIleGluGluGlyTyrGluProasp 70  
DB 259 GAGCCGATTGGCCCAAGTGTTCCTTTCCTTAAGAAATTGGAAGCTGGGAACCCGAT 318

QY 71 Asp 71  
DB 319 GAC 321

RESULT 25  
AB013819  
LOCUS AB013819 955 bp mRNA linear ROD 25-FEB-1999  
DEFINITION Mus musculus mRNA for TIAP, complete cds.  
ACCESSION AB013819  
VERSION AB013819.1 GI:3135206  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (sites)  
Kobayashi,K., Hatano,M., Otaki,M., Ogasawara,T. and Tokuhisa,T.  
Expression of a murine homologue of the inhibitor of apoptosis  
protein is related to cell proliferation  
Proc. Natl. Acad. Sci. U.S.A. 96 (4), 1457-1462 (1999)  
MEDLINE 99145571  
PUBMED 9990045

REFERENCE 2 (bases 1 to 955)  
AUTHORS Kobayashi,K., Otaki,M., Ogasawara,T. and Tokuhisa,T.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAY-1998) Koichi Kobayashi, Chiba University Graduate  
School of Medicine, Department of Developmental Genetics, Inohana  
1-8-1, Chuo-ku, Chiba 260-8670, Japan  
(E-mail:kobayashi@devgenol.m.chiba-u.ac.jp, Tel.81-43-226-2182)  
COMMENT  
FEATURES  
source 1..955  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="embryo"  
/dev\_stage="day16"  
1..955  
/gene="tiap"  
/gene="tiap"  
109..531  
/note="IAP repeat"  
/codon\_start=1  
/product="TIAP"  
/protein\_id="BAA28266.1"  
/db\_xref="GI:3135207"  
/translation="MGAPALPOLIOLYLNRYIAFKWMPLEDCATPERMAAGFI  
HCPTEHPDLAOCFFCFKELBEMEPDNDIEBHRKHSRSCAFITVKOMEELTVSEFL  
KIDROPARKKIATKNTNNKKEFEETAKTTRQSI EQLAA"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.5e-38 Length: 955  
Score: 41.00 Matches: 41  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 28.87% Indels: 0  
DB: 10 Gaps: 0

US-09-690-825-34 (1-142) x AB013819 (1-955)

QY 31 CysAlaCysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrgluAsn 50  
DB 199 TGGCGCTGCACCCGAGAGGAATGGCGAGGCTGCTTCATCCTCACTGCCCTACCGAGAAC 258  
QY 51 GluProAspLeuAlaGlnCysPhePheCysPheIleGluGluGlyTyrGluProasp 70  
DB 259 GAGCCGATTGGCCCAAGTGTTCCTTTCCTTAAGAAATTGGAAGCTGGGAACCCGAT 318  
QY 71 Asp 71  
DB 319 GAC 321

RESULT 26  
 LOCUS BC004702 3352 bp mRNA linear ROD 03-OCT-2003  
 DEFINITION Mus musculus baculoviral IAP repeat-containing 5, mRNA (cDNA clone MGC:5790 IMAGE:3499203), complete cds.  
 ACCESSION BC004702  
 VERSION BC004702.1 GI:13435665  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 3352)  
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, K.H., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Locellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McErmann, K.J., Malek, O.A., Gunaratne, P.H., Richards, S., Wooley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shavchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butler, L.A., S., Kizylinski, M.I., Skalska, U., Smallos, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Warr, M.A.,  
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 22388257  
 12477932  
 2 (bases 1 to 3352)  
 Strausberg, R.  
 Direct Submission  
 Submitted (21-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2530, USA  
 REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov  
 Contact: MGC help desk  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Jeffrey Green M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
 Center code: BCM-HGSC  
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
 Contact: amg@bcm.tmc.edu  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H., Kovits, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanaavati, A.N., Gibbs, R.A.  
 FEATURES  
 Source  
 1..3352  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="MGC:5790 IMAGE:3499203"  
 /tissue\_type="mammary tumor; C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."  
 /clone\_lib="NCI\_CGAP\_Mam6"

gene  
 1..3352  
 /gene="Birc5"  
 /note="synonyms: TIAP, AAC-11"  
 /db\_xref="LocustID:11799"  
 /db\_xref="MGI:1203517"  
 33..398  
 /codon\_start=1  
 /product="Birc5 protein"  
 /protein\_id="AAH04702.1"  
 /db\_xref="GI:13435666"  
 /db\_xref="LocustID:11799"  
 /translation="MGAPALPOLWOLYLNKRYATPKMPPLEDCACTPERMAAGFI  
 HCPTEHPPLAQCFPCPKLESGEPEDNPRIEHRKHSQCAFLTVKQMEELVSEFL  
 KIDRORAKKIVCMLEND"  
 72..293  
 /note="BIR; Region: Baculoviral inhibition of apoptosis protein repeat"  
 /db\_xref="CD:smart00238"  
 misc\_feature  
 Alignment Scores:  
 Pred. No.: 1,2e-37 Length: 3352  
 Score: 41.00 Matches: 41  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 28.87% Indels: 0  
 DB: 10 Gaps: 0  
 US-09-690-825-34 (1-142) x BC004702 (1-3352)  
 QY 31 CysAlaCysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrGluAsn 50  
 Db 123 TGGCGCTGCACCCGACGAGGAATGGCGGAGGCTGCTCATCGACTGCCCTACCGAGAAC 182  
 QY 51 GluProAspLeuAlaGlnCysPhePheCysPheLeuGluGlnGlyTyrGluProAsp 70  
 Db 183 GAGCCCTGATTGGCCGACGATTTTCTTCCTTAAGAAATGGAAGCTGGAGACCCGAT 242  
 QY 71 Asp 71  
 Db 243 GAC 245  
 RESULT 27  
 AF276775 903 bp mRNA linear ROD 12-JUL-2000  
 LOCUS AF276775  
 DEFINITION Rattus norvegicus survivin mRNA, complete cds.  
 ACCESSION AF276775  
 VERSION AF276775.1 GI:5050024  
 KEYWORDS  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 1 (bases 1 to 903)  
 Chen, D., Cao, G., and Chen, J.  
 Molecular cloning and characterization of rat survivin  
 Unpublished  
 2 (bases 1 to 903)  
 Chen, D., Cao, G., and Chen, J.  
 Direct Submission  
 Submitted (09-JUN-2000) Neurology, University of Pittsburgh, 3500 Terrace Street, Pittsburgh, PA 15261, USA  
 Location/Qualifiers  
 1..903  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 5..433  
 /codon\_start=1

```

/product="survivin"
/protein_id="AAF82586.1"
/db_xref="GI:9050025"
/translation="MGATLAPPIQMOMVKDRIYFNKWPPELEDCSPERRMAGFI
HPTENEPDLACCFCKEKLBEEMEDDNPFEHHRHSGCAPLTVKKQVEELTVSEFL
KIDKQAKKIAKETNNKQKEFEETRTVROSIEQLALR"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 7,48e-36 Length: 903
Score: 39.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.46% Indels: 0
DB: 10 Gaps: 0

```

US-09-690-825-34 (1-142) x AF276775 (1-903)

```

QY 33 CysThrProGluArgMetLacGluPheHleHisCysProThrGluAspGluPro 52
DB 101 TGCACCCCGACGAGCGGAGGCTGGCTTCATCCAGTCCGACGAGATGAGCTT 160

```

```

QY 53 AspLeuAlaGlnCysPhePheCysPheLysGluLeuGluGlyTrpGluProAspAsp 71
DB 161 GATTTCGCCAGTGTTTCTTCGCTTTAAGAACTGGAAGCTGGGAACCGGATGAC 217

```

```

RESULT 28
AR157507 AR157507 14796 bp DNA linear PAT 17-OCT-2001
LOCUS AR157507 Sequence 35 from patent US 6245523.
DEFINITION AR157507
ACCESSION AR157507
VERSION AR157507.1 GI:16218452
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 14796)
AUTHORS Altieri,D.C.
TITLE Survivin, a protein that inhibits cellular apoptosis, and its
JOURNAL modulation Patent: US 6245523-A 35 12-JUN-2001;
FEATURES Location/Qualifiers
source 1..14796
/mol_type="unassigned DNA"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 1.17e-34 Length: 14796
Score: 39.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.46% Indels: 0
DB: 6 Gaps: 0

```

US-09-690-825-34 (1-142) x AR157507 (1-14796)

```

QY 75 GluGluHisLysLysHisSerSerGlyCysAlaPheLeuSerValLysLysGlnPheGlu 94
DB 5159 GAGGAACATATAAAGCATTCGTCGGCTTCGCTTCCTTCGTCAAGAGAGATTGAA 5218

```

```

QY 95 GluLeuThrLeuGluGluPheLeuLysLeuAspArgGluArgAlaLysAsnLysIle 113
DB 5219 GAATTAAACCTTGTTGAATTTTGAACAGCAAGAAAGCCCAACCAAAATT 5275

```

```

RESULT 29
AR181541 AR181541 14796 bp DNA linear PAT 20-APR-2002
LOCUS AR181541 Sequence 3 from patent US 6335194.
DEFINITION AR181541
ACCESSION AR181541
VERSION AR181541.1 GI:20223755
KEYWORDS
SOURCE Unknown.

```

## ORGANISM

Unknown.

REFERENCE 1 (bases 1 to 14796)

AUTHORS Bennett,C.Frank., Ackermann,E.J., Swayze,E.B. and Cowseert,L.M.

## TITLE

Antisense modulation of survivin expression

## JOURNAL

Patent: US 6335194-A 3 01-JAN-2002;

## FEATURES

Location/Qualifiers

## SOURCE

1..14796

/organism="unknown"

/mol\_type="unassigned DNA"

## ORIGIN

Unknown.

## Alignment Scores:

```

Pred. No.: 1.17e-34 Length: 14796
Score: 39.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.46% Indels: 0
DB: 6 Gaps: 0

```

US-09-690-825-34 (1-142) x AR181541 (1-14796)

```

QY 75 GluGluHisLysLysHisSerSerGlyCysAlaPheLeuSerValLysLysGlnPheGlu 94
DB 5159 GAGGAACATATAAAGCATTCGTCGGCTTCGCTTCCTTCGTCAAGAGAGATTGAA 5218

```

```

QY 95 GluLeuThrLeuGluGluPheLeuLysLeuAspArgGluArgAlaLysAsnLysIle 113
DB 5219 GAATTAAACCTTGTTGAATTTTGAACAGCAAGAAAGCCCAACCAAAATT 5275

```

```

RESULT 30
AX333491 AX333491 14796 bp DNA linear PAT 09-JAN-2002
LOCUS AX333491 Sequence 4000 from Patent WO0194629.
DEFINITION AX333491
ACCESSION AX333491
VERSION AX333491.1 GI:18124125
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

REFERENCE 1 (bases 1 to 14796)

AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ehnert,R., Endress,G.,

## TITLE

Cancer gene determination and therapeutic screening using signature

## JOURNAL

Patent: WO 0194629-A 4000 13-DEC-2001;

## FEATURES

Location/Qualifiers

## SOURCE

1..14796

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

## ORIGIN

Unknown.

## Alignment Scores:

```

Pred. No.: 1.17e-34 Length: 14796
Score: 39.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.46% Indels: 0
DB: 6 Gaps: 0

```

US-09-690-825-34 (1-142) x AX333491 (1-14796)

```

QY 75 GluGluHisLysLysHisSerSerGlyCysAlaPheLeuSerValLysLysGlnPheGlu 94
DB 5159 GAGGAACATATAAAGCATTCGTCGGCTTCGCTTCCTTCGTCAAGAGAGATTGAA 5218

```

```

QY 95 GluLeuThrLeuGluGluPheLeuLysLeuAspArgGluArgAlaLysAsnLysIle 113
DB 5219 GAATTAAACCTTGTTGAATTTTGAACAGCAAGAAAGCCCAACCAAAATT 5275

```

RESULT 31  
AX334154 14796 bp DNA linear PAT 09-JAN-2002  
LOCUS AX334154  
DEFINITION Sequence 4663 from Patent WO0194629.  
ACCESSION AX334154  
VERSION AX334154.1 GI:18124873  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrigan, S., Soppet, D.R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature gene sets  
JOURNAL Patent: WO 0194629-A 4663 13-DEC-2001;  
FEATURES Location/Qualifiers  
1..14796  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.17e-34 Length: 14796  
Score: 39.00 Matches: 39  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 27.46% Indels: 0  
DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x AX334154 (1-14796)

QY 75 GUGUHHISLYSHISSEISERGLYCYSALAPHEUSERYALYSLYSGINPHEGLU 94  
DB 5159 GAGGAAACATTAAGACATTCGCGGTTCCGCTTCTTCGTCAAGACAGCTTGAA 5218

QY 95 GUGUHHISLYSHISSEISERGLYCYSALAPHEUSERYALYSLYSGINPHEGLU 113  
DB 5219 GAATTAACCTTGGAATTTTGAACCTGACAGAGAAAGCCAGAACAAATTT 5275

RESULT 32  
AX336483 14796 bp DNA linear PAT 09-JAN-2002  
LOCUS AX336483  
DEFINITION Sequence 6992 from Patent WO0194629.  
ACCESSION AX336483  
VERSION AX336483.1 GI:18127202  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrigan, S., Soppet, D.R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature gene sets  
JOURNAL Patent: WO 0194629-A 6992 13-DEC-2001;  
FEATURES Location/Qualifiers  
1..14796  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.17e-34 Length: 14796  
Score: 39.00 Matches: 39  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 27.46% Indels: 0  
DB: 6 Gaps: 0

Query Match: 27.46% Indels: 0  
DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x AX336483 (1-14796)

QY 75 GUGUHHISLYSHISSEISERGLYCYSALAPHEUSERYALYSLYSGINPHEGLU 94  
DB 5159 GAGGAAACATTAAGACATTCGCGGTTCCGCTTCTTCGTCAAGACAGCTTGAA 5218

QY 95 GUGUHHISLYSHISSEISERGLYCYSALAPHEUSERYALYSLYSGINPHEGLU 113  
DB 5219 GAATTAACCTTGGAATTTTGAACCTGACAGAGAAAGCCAGAACAAATTT 5275

RESULT 33  
AX410775 14796 bp DNA linear PAT 14-JUN-2002  
LOCUS AX410775  
DEFINITION Sequence 3422 from Patent WO0229103.  
ACCESSION AX410775  
VERSION AX410775.1 GI:21443480  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.  
TITLE Gene expression profiles in liver cancer  
JOURNAL Patent: WO 0229103-A 3422 11-APR-2002;  
GENE LOGIC INC (US)  
FEATURES Location/Qualifiers  
1..14796  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="EMBL/Genbank Accession No. U75285"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.17e-34 Length: 14796  
Score: 39.00 Matches: 39  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 27.46% Indels: 0  
DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x AX410775 (1-14796)

QY 75 GUGUHHISLYSHISSEISERGLYCYSALAPHEUSERYALYSLYSGINPHEGLU 94  
DB 5159 GAGGAAACATTAAGACATTCGCGGTTCCGCTTCTTCGTCAAGACAGCTTGAA 5218

QY 95 GUGUHHISLYSHISSEISERGLYCYSALAPHEUSERYALYSLYSGINPHEGLU 113  
DB 5219 GAATTAACCTTGGAATTTTGAACCTGACAGAGAAAGCCAGAACAAATTT 5275

RESULT 34  
BD192864 14796 bp DNA linear PAT 17-JUL-2003  
LOCUS BD192864  
DEFINITION Survivin, a protein that inhibits cellular apoptosis, and its modulation.  
ACCESSION BD192864  
VERSION BD192864.1 GI:33002603  
KEYWORDS UP 2002514060-A/6.  
SOURCE Cuphea hookeriana  
ORGANISM Cuphea hookeriana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Myrtales; Lythraceae; Cuphea.

REFERENCE  
AUTHORS Altieri, D.C.  
TITLE Survivin, a protein that inhibits cellular apoptosis, and its  
JOURNAL Patent: UP 2002514060-A 6 14-MAY-2002;  
YALE UNIVERSITY

**AUTHORS** Altieri, D.C., Adida, C. and Ambrosini, G.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (17-OCT-1996) Pathology, Boyer Center for Molecular Medicine Rm.436B, Yale University School of Medicine, 255 Congress

REFERENCE 2 (DASES 1 CO 134840)

REFERENCE 2 (DASES 1 CO 134840)



```

TITLE      Direct Submission
JOURNAL    Submitted (15-JAN-2003) Whitehead Institute/MIT Center for Genome
REFERENCE  Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS    3 (bases 1 to 209751)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karakas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrum,J., Menus,L., Mhova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schner,S., Schupbak,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE      Direct Submission
JOURNAL    Submitted (16-JAN-2003) Whitehead Institute/MIT Center for Genome
REFERENCE  Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS    4 (bases 1 to 209751)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Cornu,B., Erickson,K.,
Diaz,J.S., Dodge,S., Doolley,K., Dorris,L., Eriksson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karakas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Liu,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Menus,L., Mhova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schner,S., Schupbak,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

COMMENT    Direct Submission
Submitted (05-FEB-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 20, 2002 this sequence version replaced gi:26190573.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE      Genome Center
JOURNAL    Center: Whitehead Institute/ MIT Center for Genome Research
COMMENT    Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

TITLE      Project Information
JOURNAL    Center project name: L11992
COMMENT    Center clone name: 219_G_17

FEATURES   Location/Qualifiers
source     1..209751
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="17"
            /map="17"
            /clone="RP11-219G17"
            /clone_11b="RP11-219G17 Human Male BAC"
            /repeat_region
            /rpt_family="C-rich"
            /rpt_family="C-rich"

```

```

repeat_region complement(634..972)
repeat_region /rpt_family="AluX"
repeat_region complement(1653..1760)
repeat_region /rpt_family="MER3A"
repeat_region complement(1807..1886)
repeat_region /rpt_family="MIR"
repeat_region complement(2001..2312)
repeat_region /rpt_family="AluDo"
repeat_region 2336..2636
repeat_region /rpt_family="AluSx"
repeat_region complement(3684..3706)
repeat_region /rpt_family="AluDo"
repeat_region 3707..3993
repeat_region /rpt_family="AluSx"
repeat_region complement(3994..4267)
repeat_region /rpt_family="AluDo"
repeat_region 5194..5371
repeat_region /rpt_family="AluSg/X"
repeat_region 5606..5738
repeat_region /rpt_family="FRAM"
repeat_region complement(5739..5850)
repeat_region /rpt_family="Alu"
repeat_region complement(5865..5999)
repeat_region /rpt_family="AluSg/X"
repeat_region 6293..6600
repeat_region /rpt_family="AluSP"
repeat_region complement(7562..7879)
repeat_region /rpt_family="AluY"
repeat_region 8178..8326
repeat_region /rpt_family="L1ME4a"
repeat_region 8327..8617
repeat_region /rpt_family="AluSx"
repeat_region 8618..8721
repeat_region /rpt_family="L1ME4a"
repeat_region 8799..9231
repeat_region /rpt_family="L13"
repeat_region 9248..9488
repeat_region /rpt_family="L13"
repeat_region 9986..10097
repeat_region /rpt_family="AluSg/X"
repeat_region complement(10100..10397)
repeat_region /rpt_family="AluY"
repeat_region 10159..10164
repeat_region /note="<30 qual SNGL region"
repeat_region 10399..10579
repeat_region /rpt_family="AluSg/X"
repeat_region 10580..10908
repeat_region /rpt_family="AluSx"
repeat_region 10909..10919
repeat_region /rpt_family="AluSg/X"
repeat_region 11385..11522
repeat_region /rpt_family="MIR"
repeat_region complement(11696..11997)
repeat_region /rpt_family="AluY"
repeat_region complement(12826..12967)
repeat_region /rpt_family="L1MC4"
repeat_region complement(12968..13262)
repeat_region /rpt_family="AluYb"
repeat_region complement(13263..13344)
repeat_region /rpt_family="L1MC4"
repeat_region complement(13354..13532)
repeat_region /rpt_family="AluYb"
repeat_region complement(13533..13820)
repeat_region /rpt_family="AluSx"
repeat_region complement(13821..13890)
repeat_region /rpt_family="L1MC4"
repeat_region 14283..14348
repeat_region /rpt_family="MER45C"
repeat_region complement(14349..14525)
repeat_region /rpt_family="AluSg"
repeat_region complement(14526..14828)
repeat_region /rpt_family="AluY"
repeat_region complement(14829..14962)

```



```

repeat_region      /rpt_family="Alusg"
                    14963..14999
repeat_region      /rpt_family="MER45C"
                    15000..15299
repeat_region      /rpt_family="Alusx"
                    15303..15598
repeat_region      /rpt_family="Alusx"
                    15599..15752
repeat_region      /rpt_family="MER45C"
                    complement(15198..16332)
repeat_region      /rpt_family="WIR"
                    16832..16961
repeat_region      /rpt_family="WIR"
                    17716..17810
unrepeated          /note="single clone coverage"
repeat_region      17794..17975
repeat_region      /rpt_family="CT-rich"
                    18235..18553
repeat_region      /rpt_family="Alusg"

```

Alignment Scores:

Pred. No.:	1.59e-33	Length:	209751
Score:	39.00	Matches:	39
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	27.46%	Indels:	0
DB:	9	Gaps:	0

US-09-690-825-34 (1-142) x AC087645 (1-209751)

QY 75 GluGluHisLysLysHisSerSerGlyCysAlaPheLeuSerValLysLysGlnPheGlu 94  
 DB 125460 GAGGACATATAAAGCATTCGCGCGTTCGCGCTTCCTTCGTCAGAGACAGCTTTGAA 125401

QY 95 GluLeuThrLysGluPheLeuLysLeuSpArgGluArgAlaLysAsnLysIle 113  
 DB 125400 GAATTACCCCTTGGAATTGTAAGTGGACAGAAAGGCCAGACAAATTT 125344

RESULT 38  
 ARI84472 417 bp DNA linear PAT 20-APR-2002  
 LOCUS ARI84472  
 DEFINITION Sequence 1 from patent US 6346389.  
 ACCESSION ARI84472.1 GI:20230437  
 VERSION ARI84472.1  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 417)  
 AUTHORS Aliteri,D.C.  
 TITLE Method for selectively modulating the interactions between survivin  
 and tubulin  
 JOURNAL Patent: US 6346389-A 1 12-FEB-2002;  
 FEATURES Location/Qualifiers  
 source 1..417  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.9e-34 Length: 417  
 Score: 37.00 Matches: 37  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 26.06% Indels: 0  
 DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x ARI84472 (1-417)  
 QY 1 MetGlyAlaProThrLeuProPheLeuGluGlyCysAlaCysThrProGluArg 20  
 DB 291 ATGGGAGCCCGACGCTTGGCCCTGCGCTTGGACGCGCTTCTCAAGGACCCGCACTTCT 350

QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArg 37  
 DB 351 ACATTCAGAACTGGCCCTTCTTGAAGGCTGCGCTGCTCACCCCGACGCG 401

RESULT 39  
 AC032035 154840 bp DNA linear HTG 08-SEP-2000  
 LOCUS AC032035  
 DEFINITION Homo sapiens chromosome 17 clone RP11-141D15 map 17, \*\*\* SEQUENCING  
 IN PROGRESS \*\*\*, 18 unordered pieces.  
 ACCESSION AC032035.3 GI:9994161  
 VERSION AC032035  
 KEYWORDS HTG; HTGS; PHASE1.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 154840)  
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 TITLE Homo sapiens chromosome 17, clone RP11-141D15  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 154840)  
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Adnan,H., Allen,N.,  
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
 Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G., F.,  
 Campolino,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
 Collymore,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S.,  
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,  
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
 Grand-pierre,N., Grant,G., Hagos,B., Harford,A., Horton,L.,  
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karakas,A.,  
 Klein,U., Lacroque,K., Lamazares,R., Landers,T., Lechoczky,J.,  
 Levine,R., Lien,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
 McCarty,M., McEwan,P., McGuirk,A., McKernan,K., McNetters,R.,  
 Melchior,J., Meneses,L., Mihova,T., Miranda,C., Mleaga,V., Morrow,J.,  
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pletzer,N.,  
 Pisani,C., Pollard,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
 Stange-Thomann,N., Stojanovic,N., Subramaniam,A., Talamas,J.,  
 Testaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,  
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE  
 JOURNAL  
 COMMENT  
 Submitted (03-APR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141 USA  
 On Sep 8, 2000 this sequence version Replaced by: 7705196.  
 All repeats were identified using RepeatMasker:  
 http://ftp.genome.washington.edu/TM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L9138  
 Center clone name: 141\_D15

NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 18 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1 2373: contig of 2373 bp in length  
 \* 2374 2473: gap of 100 bp  
 \* 2474 4930: contig of 2457 bp in length  
 \* 4931 5030: gap of 100 bp  
 \* 5031 7215: contig of 2185 bp in length  
 \* 7216 7315: gap of 100 bp  
 \* 7316 9541: contig of 2226 bp in length

9542 9641: gap of 100 bp  
\* 9642 12211: contig of 2570 bp in length  
\* 12212 12311: gap of 100 bp  
\* 12312 14923: contig of 2612 bp in length  
\* 14924 15023: gap of 100 bp  
\* 15024 19094: contig of 4071 bp in length  
\* 19095 19194: gap of 100 bp  
\* 19195 22287: contig of 3092 bp in length  
\* 22287 22387: gap of 100 bp  
\* 22387 26147: contig of 3761 bp in length  
\* 26148 26247: gap of 100 bp  
\* 26248 31933: contig of 5566 bp in length  
\* 31934 32033: gap of 100 bp  
\* 32034 37339: contig of 5306 bp in length  
\* 37340 37440: gap of 100 bp  
\* 37440 42858: contig of 5419 bp in length  
\* 42859 42959: gap of 100 bp  
\* 42959 48810: contig of 5852 bp in length  
\* 48811 48910: gap of 100 bp  
\* 48911 56029: contig of 7119 bp in length  
\* 56030 56129: gap of 100 bp  
\* 56130 62189: contig of 6060 bp in length  
\* 62190 62289: gap of 100 bp  
\* 62290 82959: contig of 20670 bp in length  
\* 82960 83059: gap of 100 bp  
\* 83060 107059: contig of 24000 bp in length  
\* 107060 107159: gap of 100 bp  
\* 107160 154840: contig of 47681 bp in length.  
Location/Qualifiers  
1. 154840  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17p"  
/clone="RP11-141D15"  
/clone\_lib="RPC1-11 Human Male BAC"

ORIGIN

Alignment Scores:  
Pred. No.: 2 666-31 length: 154840  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 26.06% Indels: 0  
DB: 2 Gaps: 0

US-09-690-825-34 (1-142) x AC032035 (1-154840)

QY 1 MetGlyAlaProThrLeuProAlaTyrGlnProPheLeuLysAspHisArgIleSer 20

Db 31031 ATGGGGGCCCCGACGTTGGCCCTGCTGGCAGCCCTTTCTCAAGACACCGCATCTCT 31090

QY 21 ThrPheLysAsnTyrProPheLeuGluGlyCysAlaCysThrProGluArg 37

Db 31091 ACATTCAAGAACTGGCCCTTTTGAAGGGCTGCGCTGCACCCCGGACGG 31141

RESULT 40

AC010532 229426 bp DNA linear HTG 05-SEP-2000

AC010532 Homo sapiens chromosome 17 clone RP11-219G17, LOW-PASS SEQUENCE

LOCUS SAMPLING.

AC010532

VERSION AC010532.2 GI:6758797

KEYWORDS HTG; HTGS\_PHASE0.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 229426)

DOE Joint Genome Institute.

TITLE Sequencing of Human Chromosome 17

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 229426)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Jan 26, 2000 this sequence version replaced gi:5882406.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----

\* NOTE: This record contains 99 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 563: contig of 563 bp in length  
\* 564 1436: gap of unknown length  
\* 1437 1649: gap of 213 bp in length  
\* 1650 2523: gap of unknown length  
\* 2524 3286: gap of unknown length  
\* 3287 3613: gap of 327 bp in length  
\* 3614 5430: gap of unknown length  
\* 5431 6681: gap of 1251 bp in length  
\* 6682 7357: gap of 676 bp in length  
\* 7358 8117: gap of 760 bp in length  
\* 8118 8718: gap of unknown length  
\* 8719 9464: gap of 746 bp in length  
\* 9465 9964: gap of 500 bp in length  
\* 9965 10959: gap of 995 bp in length  
\* 10960 11873: gap of 914 bp in length  
\* 11874 12656: gap of 783 bp in length  
\* 12657 13674: gap of unknown length  
\* 13675 14935: gap of 1261 bp in length  
\* 14936 16423: gap of unknown length  
\* 16424 17186: gap of 763 bp in length  
\* 17187 18079: gap of unknown length  
\* 18080 18813: gap of 734 bp in length  
\* 18814 19729: gap of 916 bp in length  
\* 19730 20675: gap of 946 bp in length  
\* 20676 21721: gap of unknown length  
\* 21722 23111: gap of 1390 bp in length

```

* 23112 25188: contig of 2077 bp in length
* 25189 26588: contig of 1400 bp in length
* 26589 27644: contig of 1056 bp in length
* 27645 29372: contig of 1728 bp in length
* 29373 29495: contig of 123 bp in length
* 29496 30431: contig of 936 bp in length
* 30432 31449: contig of 1018 bp in length
* 31450 32983: contig of 1534 bp in length
* 32984 33162: contig of 179 bp in length
* 33163 33975: contig of 813 bp in length
* 33976 35031: contig of 1056 bp in length
* 35032 35709: contig of 678 bp in length
* 35710 37308: contig of 1599 bp in length
* 37309 38107: contig of 799 bp in length
* 38108 39102: contig of 995 bp in length
* 39103 39591: contig of 489 bp in length
* 39592 41151: contig of 1560 bp in length
* 41152 42842: contig of 1691 bp in length
* 42843 43828: contig of 986 bp in length
* 43829 44009: contig of 181 bp in length
* 44010 44874: contig of 865 bp in length
* 44875 45191: contig of 317 bp in length
* 45192 46147: contig of 956 bp in length
* 46148 47504: contig of 1357 bp in length
* 47505 49091: contig of 1587 bp in length
* 49092 50259: contig of 1168 bp in length
* 50260 51361: contig of 1102 bp in length
* 51362 52631: contig of 1270 bp in length
* 52632 53607: contig of 976 bp in length
* 53608 54820: contig of 1213 bp in length
* 54821 56600: contig of 1780 bp in length
* 56601 58236: contig of 1636 bp in length
* 58237 59716: contig of 1480 bp in length
* 59717 62221: contig of 2505 bp in length
* 62222 65085: contig of 2864 bp in length
* 65086 66739: contig of 1654 bp in length
* 66739: contig of 1654 bp in length
* 66739: gap of unknown length

```

```

* 66740 68359: contig of 1620 bp in length
* 68360 70014: contig of 1655 bp in length
* 70015 71808: contig of 1794 bp in length
* 71809 73581: contig of 1773 bp in length
* 73582 75574: contig of 1993 bp in length
* 75575 78468: contig of 2894 bp in length
* 78469 79231: contig of 763 bp in length
* 79232 83240: contig of 4009 bp in length
* 83241 86157: contig of 2917 bp in length
* 86158 89150: contig of 2993 bp in length
* 89151 92671: contig of 3521 bp in length
* 92672 95442: contig of 2771 bp in length
* 95443 98476: contig of 3034 bp in length
* 98477 100815: contig of 2339 bp in length
* 100816 103005: contig of 2190 bp in length
* 103006 105514: contig of 2509 bp in length
* 105515 108704: contig of 3190 bp in length
* 108705 111097: contig of 2393 bp in length
* 11098 114888: contig of 3791 bp in length
* 114889 117423: contig of 2535 bp in length
* 117423: gap of unknown length

```

## Alignment Scores:

```

Pred. No.: 3,92e-31 Length: 229426
Score: 37.00 Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.06% Indels: 0
DB: 2 Gaps: 0

```

US-09-690-825-34 (1-142) x ACO10532 (1-229426)

```

QY 1 MetGjAlaProhrrhneupProaIATpGInProPhelenuIsasphIsaTgIleser 20
Db 77692 ATGGGTCGCCGAGCGTGGCCCTGCTGCGACCCCTTCTCAAGACACCCGCACTCT 77751
QY 21 ThrPhelystAntIrrProPhelenuGjngIyCsAlaCysThrProGluArg 37
Db 77752 ACAATTCAAGAACTGGCCCTTCTTGAGGGGCTGGCGCTGACCCCGAGCGG 77802

```

```

RESULT 41
CRUCIRA 779 bp DNA linear ROD 27-APR-1993
LOCUS Cricetus griseus carcinogen-induced rearranged DNA.
DEFINITION M80243
ACCESSION M80243
VERSION GI:191034
SOURCE Cricetus griseus (Chinese hamster)
ORGANISM Cricetus griseus Chordata: Craniata: Vertebrata: Euteleostomi:
Eukaryota: Metazoa: Mammalia: Rodentia: Sciurognathi: Muridae: Cricetinae:
Cricetus.
1 (bases 1 to 779)
REFERENCE
AUTHORS Barr,F.G., Davis,R.J., Eichenfield,L. and Emanuel,B.S.

```

TITLE Structural analysis of a carcinogen-induced genomic rearrangement  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (3), 942-946 (1992)  
 MEDLINE 92141234  
 PUBMED 1736310  
 COMMENT Original source text: Cricetus griseus (Library: Lambda Charon 35) DNA.

FEATURES  
 source Location/Qualifiers  
 1..779  
 /organism="Cricetus griseus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10029"  
 /haplotype="---"  
 /cell\_line="RJK92 - V79 derivative"  
 /cell\_type="fibroblast"  
 /tissue\_id="Lambda Charon 35"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,96e-30 Length: 779  
 Score: 34.00 Matches: 34  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 23.94% Indels: 0  
 DB: 10 Gaps: 0

US-09-690-825-34 (1-142) x CRUCIRA (1-779)

QY 38 MetAlaGUAaGlyPheIleHisCysProthrgluasngluProaspLeuAlaGlnCys 57

DB 581 ATGGCGAAGCGCGCTTCACTCCCTCCCTCCGAAACGAGCTGACTGGCCAGTGT 522

QY 58 PhePheCysPhePheGluGluGluGlyTTPGluProaspAsp 71

DB 521 TTTTCTGCTTTAAGAGCTGGAAGCTGGAGCAGATGAC 480

RESULT 42  
 AF077351 9263 bp DNA linear ROD 20-DEC-2000

LOCUS Mus musculus inhibitor of apoptosis homolog gene, complete cds.

ACCESSION AF077351 GI:11907561

VERSION AF077351.1

KEYWORDS Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Uren,A.G., Wong,L., Pakusch,M., Fowler,K.J., Burrows,F.J.,  
 Vaux,D.L. and Choo,K.H.

TITLE Survivin and the inner centromere protein INCENP show similar  
 cell-cycle localization and gene knockout phenotype

REFERENCE Curr. Biol. 10 (21), 1319-1326 (2000)

AUTHORS Uren,A.U. and Vaux,D.L.

TITLE Submitted (08-JUN-1998) Molecular Cancer Division, The Walter and  
 Eliza Hall Institute, Royal Parade, Parkville, Vic 3052, Australia

FEATURES  
 source Location/Qualifiers  
 1..9263

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="Black 6"

/db\_xref="taxon:10090"

/product="inhibitor of apoptosis homolog"

/note="TAP"

/codon\_start=1

/protein\_id="AAG41214.1"

/db\_xref="GI:11907562"  
 /translation="MGAPALPOLWOLYKNYRIATPKNMPLEDCATPERMAAGFI  
 HCPTEHPDLACFCFCFKELBGMEDDNDPIEHRKSGSCAFITVKKOMEITVSEFL  
 KIDRQARKXIKAKETNNKQKEFEETAKTTROSIEQLAA"

ORIGIN

Alignment Scores:  
 Pred. No.: 5.67e-29 Length: 9263  
 Score: 34.00 Matches: 34  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 23.94% Indels: 0  
 DB: 10 Gaps: 0

US-09-690-825-34 (1-142) x AF077351 (1-9263)

QY 38 MetAlaGUAaGlyPheIleHisCysProthrgluasngluProaspLeuAlaGlnCys 57

DB 2320 ATGGCGAAGCGCGCTTCACTCCCTCCCTCCGAAACGAGCTGACTGGCCAGTGT 2379

QY 58 PhePheCysPhePheGluGluGluGlyTTPGluProaspAsp 71

DB 2380 TTTTCTGCTTTAAGAGATGGAAGCTGGAAACCGATGAC 2421

RESULT 43

AB036741 10074 bp DNA linear ROD 22-DEC-2000

LOCUS Mus musculus gene for TAP, complete cds.

ACCESSION AB036741 GI:11990425

VERSION AB036741.1

KEYWORDS TAP.

ORGANISM Mus musculus (house mouse)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Otake,M., Hatano,M., Kobayashi,K., Ogasawara,T., Kuriyama,T. and  
 Tokuhisa,T.

TITLE Cell cycle dependent regulation of TAP/m-survivin expression

REFERENCE Unpublished

AUTHORS Otake,M., Hatano,M. and Tokuhisa,T.

TITLE Submitted (07-JAN-2000) Masayuki Otake, Chiba University Graduate  
 School of Medicine, Department of Developmental Genetics, 1-8-1  
 Inohana Chuo-ku, Chiba, Chiba 260-8670, Japan

URL:http://devgenol.m.chiba-u.ac.jp/, Tel:81-43-226-2182,  
 Fax:81-43-226-2183

FEATURES  
 source Location/Qualifiers  
 1..10074

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="129/Sv"

/db\_xref="taxon:10090"

/product="TAP"

/note="TAP"

/codon\_start=1

/protein\_id="BAB19787.1"

/translation="MGAPALPOLWOLYKNYRIATPKNMPLEDCATPERMAAGFI  
 HCPTEHPDLACFCFCFKELBGMEDDNDPIEHRKSGSCAFITVKKOMEITVSEFL  
 KIDRQARKXIKAKETNNKQKEFEETAKTTROSIEQLAA"

/db\_xref="GI:11907562"  
 /translation="MGAPALPOLWOLYKNYRIATPKNMPLEDCATPERMAAGFI  
 HCPTEHPDLACFCFCFKELBGMEDDNDPIEHRKSGSCAFITVKKOMEITVSEFL  
 KIDRQARKXIKAKETNNKQKEFEETAKTTROSIEQLAA"

ORIGIN

Alignment Scores:  
 Pred. No.: 6.15e-29 Length: 10074  
 Score: 34.00 Matches: 34  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0

US-09-690-825-34 (1-142) x AF077351 (1-9263)

QY 38 MetAlaGUAaGlyPheIleHisCysProthrgluasngluProaspLeuAlaGlnCys 57

DB 2320 ATGGCGAAGCGCGCTTCACTCCCTCCCTCCGAAACGAGCTGACTGGCCAGTGT 2379

QY 58 PhePheCysPhePheGluGluGluGlyTTPGluProaspAsp 71

DB 2380 TTTTCTGCTTTAAGAGATGGAAGCTGGAAACCGATGAC 2421

RESULT 43

AB036741 10074 bp DNA linear ROD 22-DEC-2000

LOCUS Mus musculus gene for TAP, complete cds.

ACCESSION AB036741 GI:11990425

VERSION AB036741.1

KEYWORDS TAP.

ORGANISM Mus musculus (house mouse)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Otake,M., Hatano,M., Kobayashi,K., Ogasawara,T., Kuriyama,T. and  
 Tokuhisa,T.

TITLE Cell cycle dependent regulation of TAP/m-survivin expression

REFERENCE Unpublished

AUTHORS Otake,M., Hatano,M. and Tokuhisa,T.

TITLE Submitted (07-JAN-2000) Masayuki Otake, Chiba University Graduate  
 School of Medicine, Department of Developmental Genetics, 1-8-1  
 Inohana Chuo-ku, Chiba, Chiba 260-8670, Japan

URL:http://devgenol.m.chiba-u.ac.jp/, Tel:81-43-226-2182,  
 Fax:81-43-226-2183

FEATURES  
 source Location/Qualifiers  
 1..10074

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="129/Sv"

/db\_xref="taxon:10090"

/product="TAP"

/note="TAP"

/codon\_start=1

/protein\_id="BAB19787.1"

/translation="MGAPALPOLWOLYKNYRIATPKNMPLEDCATPERMAAGFI  
 HCPTEHPDLACFCFCFKELBGMEDDNDPIEHRKSGSCAFITVKKOMEITVSEFL  
 KIDRQARKXIKAKETNNKQKEFEETAKTTROSIEQLAA"

Query Match: 23.94% Indels: 0  
DB: 10 Gaps: 0

US-09-690-825-34 (1-142) x AB036741 (1-10074)

QY 38 MetAlaGluAlaGlyPheIleHisCySProThrGluAsnGluProAspLeuAlaGlnCys 57  
3227 ATGGCGGAGCGCTGCTTATCCACTGCCCTACCGAGACGAGCGCTGATTGGCCCAAGTGT 3286

QY 58 PhepHeCyPheIyGluLeuGluGlyTTPGluProAspPaP 71  
3287 TTTTCTGCTTTAAGGATTTGAAGCGTGGAGCCCGATGAC 3328

RESULT 44  
AF115517 10919 bp DNA linear ROD 20-APR-1999  
LOCUS Mus musculus survivin40, survivin121, and survivin140 genes.  
DEFINITION Alternative splice products, complete cds.  
ACCESSION AF115517  
VERSION AF115517.1 GI:4588767  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 10919)  
Conway, E.M., Pollefeys, S., Corneliussen, J., Debere, I., Ong, K., Collier, D. and Schuh, A.C.  
Differential expression of three murine survivin cDNA variants  
Unpublished  
2 (bases 1 to 10919)  
Conway, E.M., Pollefeys, S., Corneliussen, J., Debere, I., Ong, K., Collier, D. and Schuh, A.C.  
Direct Submision  
Submitted (21-DEC-1998) Center for Transgene Technology, University of Leuven, Herestraat 49, 9th Floor, Leuven B-3000, Belgium  
FEATURES  
source 1. 10919  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/map="E2"  
/dev\_stage="14.5 dpc embryo"  
join(2865..2975,3268..3377,6175..6292,8741..58824)  
/product="survivin140"  
join(2865..2975,3268..3377,6175..6292,8741..8824)  
/note="member of inhibitor of apoptosis (IAP) family"  
/codon\_start=1  
/product="survivin140"  
/protein\_id="AAD26201.1"  
/db\_xref="GI:4588767"  
/translation="MGAPALPQIWQLYKRYRIATFKNWPFLDCACTPERMAAGT  
HPTENEDPLAOCFCFKSLGEMEDDNPDIHHRHSGCAFLTVKQMEELVSEPL  
KIDROAKNKIKVEMEND"  
join(2865..2975,6175..>6186)  
/product="survivin40"  
join(2865..2975,6175..6186)  
/note="resulting from removal of exon-2 derived sequences  
by alternative pre-mRNA splicing"

ORIGIN  
Alignment Scores:  
Pred. No.: 6.66e-29 Length: 10919  
Score: 34.00 Matches: 34  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 23.94% Indels: 0  
DB: 10 Gaps: 0

US-09-690-825-34 (1-142) x AF115517 (1-10919)

QY 38 MetAlaGluAlaGlyPheIleHisCySProThrGluAsnGluProAspLeuAlaGlnCys 57  
3268 ATGGCGGAGCGCTGCTTATCCACTGCCCTACCGAGACGAGCGCTGATTGGCCCAAGTGT 3327

QY 58 PhepHeCyPheIyGluLeuGluGlyTTPGluProAspPaP 71  
3328 TTTTCTGCTTTAAGGATTTGAAGCGTGGAGCCCGATGAC 3369

RESULT 45  
AL954690 197194 bp DNA linear HTG 15-NOV-2002  
LOCUS Mus musculus chromosome 11 clone RP23-386A4.  
ACCESSION AL954690  
VERSION AL954690.2 GI:25136861  
KEYWORDS HTG; HTGS PHASE2; HTGS CANCELLED.  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 197194)  
Tromans, A.  
Direct Submision  
Submitted (13-NOV-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonequest@sanger.ac.uk  
On Nov 19, 2002 this sequence version replaced gi:25045706.  
Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.  
-----Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
-----Project Information  
Center project name: bm386A4  
-----Summary Statistics  
Assembly program: XAP4, version 4.5  
Chemistry: Dye-terminator, 100% of reads  
Consensus quality: 195990 bases at least Q40  
Consensus quality: 196234 bases at least Q30  
Consensus quality: 196338 bases at least Q20  
Insert size: 197194; sum-of-contigs  
Insert size: 189681; 3.6% error; agarose-fp  
Quality coverage: 7.24x in Q20 bases; sum-of-contigs Quality  
coverage: 7.53x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.  
 1 197194: contig of 197194 bp in length.  
 Location/Qualifiers  
 source

## misc\_feature

1. 197194  
 /organsm="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="11"  
 /clone\_1lb="RP23-386A4"  
 /clone\_1lb="RP23-386A4"  
 /note="assembly fragment:02398  
 clone\_end:SP6  
 vector\_side:left"

## ORIGIN

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	1.15e-27	197194	34.00	0	0	0
Percent Similarity:	100.00%					
Best Local Similarity:	100.00%					
Query Match:	23.94%					

US-09-690-825-34 (1-142) x AL594690 (1-197194)

Qy 38 MetAlaGluAaGlyPheHleHicCysProThrGluAsnGluProAspLeuAlaGlnCys 57

Db 67096 ATGGCGAGGCTGGCTTCACTCCCTACCGAGAACGAGCTGATTTGGCCACGTGT 67155

Qy 58 PhePheCysPheYsgLueGluGluGlyTTPGluProAspAsp 71

Db 67156 TTTTCTGCTTTAAGGAATTGGAAGGCTGGGAACCGATGAC 67197

## RESULT 46

## AL591433

## LOCUS

203336 bp DNA linear ROD 21-DEC-2002  
 Mouse DNA sequence from clone RP23-268N22 on chromosome 11,  
 complete sequence.

AL591433  
 AL591433.6 GI:27368257

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Mus musculus (house mouse)  
 Mus musculus  
 Buxarjora; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 203336)  
 Bates, K.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Direct Submission  
 Submitted (20-DEC-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Dec 23, 2002 this sequence version replaced gi:25136610.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 only a small overlap as described above.  
 The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information  
 on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep  
 from the RPI-23 Mouse PAC Library  
 constructed by the group of Pieter de Jong.  
 For further details see http://www.chori.org/bacpac/home.htm  
 VECTOR: pBACg3.6  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquery@sanger.ac.uk

-----  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest, except on the rare  
 occasion of the clone being a YAC.  
 Sequence from the Mouse Genome Sequencing Consortium whole genome  
 shotgun may have been used to confirm this sequence. Sequence data  
 from the whole genome shotgun alone has only been used where it has  
 a phred quality of at least 30.  
 Location/Qualifiers  
 source

## FEATURES

1. 203336  
 /organsm="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="11"  
 /clone\_1lb="RP23-268N22"  
 /clone\_1lb="RP23-268N22"

## ORIGIN

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	1.18e-27	203336	34.00	0	0	0
Percent Similarity:	100.00%					
Best Local Similarity:	100.00%					
Query Match:	23.94%					

US-09-690-825-34 (1-142) x AL591433 (1-203336)

Qy 38 MetAlaGluAaGlyPheHleHicCysProThrGluAsnGluProAspLeuAlaGlnCys 57

Db 26097 ATGGCGAGGCTGGCTTCACTCCCTACCGAGAACGAGCTGATTTGGCCACGTGT 26156

Qy 58 PhePheCysPheYsgLueGluGluGlyTTPGluProAspAsp 71

Db 26157 TTTTCTGCTTTAAGGAATTGGAAGGCTGGGAACCGATGAC 26198

## RESULT 47

## AC143833

## LOCUS

143111 bp DNA linear HMG 09-APR-2003  
 Macaca mulatta clone CH250-268D1, \*\*\* SEQUENCING IN PROGRESS \*\*\*.

AC143833.1 GI:29649344

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Macaca mulatta (rhesus monkey)  
 Macaca mulatta  
 Buxarjora; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Macaca.

1 (bases 1 to 143111)  
 Ceaus, M. and Miosavljivic, A.

1 (bases 1 to 143111)  
 Pooled genomic indexing (PGI): mathematical analysis and experiment  
 design

## JOURNAL

## TITLE

## AUTHORS

## REFERENCE

ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI  
 2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;  
 Springer (2002)  
 2 (bases 1 to 143111)  
 Miosavljivic, A., Sodergren, E., Ceaus, M., Li, B., Jackson, A.R.,  
 Adams, C., Adio-Ogunola, B., Ali-ouman, F.R., Allen, C., Alpbach, S.L.,  
 Amarantunga, H.C., Are, J.R., Ayale, M., Banks, T., Barbieri, J.,  
 Benton, J., Biraghe, K., Blankenburg, K., Bonnin, D., Bouck, D.,  
 Bowie, S., Briteva, M., Brown, M., Brown, M., Bryant, N.P., Buhay, C.,  
 Burch, P., Burkett, C., Burdell, K.L., Byrd, N.C., Caron, T.F.,  
 Carter, M., Cavazos, S.R., Chacko, D., Chavez, D., Chen, G., Chen, R.,  
 Chen, Z., Chiu, D., Chowdhury, I., Christopoulos, C., Cleveland, C.D.,  
 Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,  
 Davy, Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,  
 Dem, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Dugan-Rocha, S., Durbin, K.J., Egan, A., Earnhart, C., Edwards, C.C., Elhai, C., Emerling, S., Escotto, M., Falls, T., Ferraruto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabrieli, A., Geo, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gottlieb, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, U., Korah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louleghed, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Marondel, I., Martin, R., Martindale, A., Martinez, B., Massey, E., Mathew, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzger, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkenwo, S., Ogun, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shin, C., Shochet, N., Sisson, I., Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Swalek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, K., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wellington, S., Williams, G., Williamson, A., Wiczcyk, R., Woodem, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Zuchelapatti, R., Weinstein, G., and Gibbs, R.

COMMENT

Center: Baylor College of Medicine  
Center code: BCM  
Web Site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
Project Information  
Center project name: LDDs  
Center clone name: CH250-268D1  
Summary Statistics  
Chemistry: Dye-terminator Big Dye; inf% of reads  
Consensus quality: 21640 bases at least Q40  
Consensus quality: 24831 bases at least Q20  
Consensus quality: 28074 bases at least Q20

FEATURES

source  
1. 143111  
/organism="Macaca mulatta"  
/mol\_type="genomic DNA"  
\* NOTE: Batted insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_drift\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html))  
\* NOTE: The contigs are based on the application  
\* of the PGI method using the Human genome (NCBI build 31)  
\* as the comparative genome.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
1 143111: contig of 143111 bp in length.  
1 Location/Qualifiers

/db\_xref="taxon:9544"  
/clone="CH250-268D1"  
1. 143111  
/note="assembly name: CH250-268D1.1B  
CONFIDENCE: 0.83"

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
2,52e-11	20.00	143111	20	0	0	0	0
Percent Similarity:	100.00%						
Best Local Similarity:	100.00%						
Query Match:	14.08%						
DB:	2						

US-09-690-825-34 (1-142) x AC142803 (1-143111)

QY 75 GUGUHUHSLYSLYSHLSERSEGYCYSAJAPHLEUSERVALLYSLYGLNPHGCU 94

DB 34047 GAGGAACATPAAAAAGCATTCACCGGTTCCGCTTCTTCTGTCAAGACACATTGCA 34106

RESULT 48

AC142802

LOCUS

DEFINITION

AC142802.1 GI:29567441

ACCESSION

VERSION

KEYWORDS

HTG; HTGS; PHASE2; HTGS\_PGI

SOURCE

Macaca mulatta

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

(in) Guigo, R. and Gusfield, D. (Eds.);  
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, MABI  
2002, ROME, ITALY, SEPTEMBER 17-21, 2002. PROCEEDINGS: 10-28;  
Springer (2002)  
2 (bases 1 to 138875)  
Milosavljevic, A., Sodergren, E., Cuiros, M., Li, B., Jackson, A.R.,  
Adams, C., Adio-Oduola, B., Ali-oshman, F.R., Allen, C., Alsbrooks, S.L.,  
Amaratunga, H.C., Are, J.R., Ayale, M., Banks, T., Barbara, D.,  
Benton, J., Binage, K., Blankenburg, K., Bonini, D., Bouck, J.,  
Bowle, S., Briviera, M., Brown, E., Brown, M., Bryant, N.P., Bunay, C.,  
Burke, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,  
Carter, M., Cavazos, S.R., Chacko, U., Chavez, D., Chen, R.,  
Chen, Z., Chiu, D., Chowdhury, I., Christopoulos, C., Cleveland, C.D.,  
Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davis, M., Davis, C.,  
Davy-Carroll, L., Dedrich, D.A., Delaney, K.R., Delgado, O.,  
Denn, A.L., Ding, Y., Dink, H.H., Douthett, K.J., Draper, H.,  
Dugan-Rocha, S., Durbin, K.J., Egan, A., Earnhart, C., Edwards, C.C.,  
Elhai, C., Emerling, S., Escotto, M., Falls, T., Ferraruto, D.,  
Flagg, N., Ford, J., Foster, P., Frantz, P., Gabrieli, A., Geo, J.,  
Garcia, A., Garner, T., Garza, N., Gill, R., Gottlieb, J.H., Guevara, W.,  
Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K.,  
Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O.,  
Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homsi, F.,  
Howard, S., Huber, J., Huliy, S., Hume, J., Ioshikhes, I., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlsson, E., Kelly, S., Khan, U., King, U., Korah, J., Kovar, C.,  
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C.,  
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,  
Louleghed, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
Ma, J., Maheshwari, M., Mapa, P., Marondel, I., Martin, R.,  
Martindale, A., Martinez, B., Massey, E., Mathew, E., McLeod, M.P.,  
Meador, M., Mei, G., Merscher, S., Metzger, M., Miller, A., Miner, G.,  
Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M.,  
Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N.,  
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkenwo, S.,  
Ogun, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,

Peery, J., Perez, L., Peters, J., Pickens, R., Primus, E., Pu, L.,  
 Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M.,  
 Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shim, C.,  
 Shooshitari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A.,  
 Stanley, H., Stone, H., Sutton, A., Svatek, A., Taylor, P., Tamerisa, A.,  
 Tameis, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B.,  
 Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D.,  
 Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R.,  
 Washington, C., Wallington, S., Williams, G., Williamson, A.,  
 Wleczky, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J.,  
 Zorrilla, S., Zuchterlapati, R., Weinstein, G. and Gibbs, R.

**TITLE**  
 Direct Submission

**JOURNAL**  
 Unpublished

**REFERENCE**  
 3 (bases 1 to 138875)

**AUTHORS**  
 Worley, K.C.

**TITLE**  
 Direct Submission

**JOURNAL**  
 Submitted (05-APR-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

**REFERENCE**  
 4 (bases 1 to 138875)

**AUTHORS**  
 Worley, K.C.

**TITLE**  
 Direct Submission

**JOURNAL**  
 Submitted (09-APR-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

**COMMENT**  
 ----- Genome Center of Medicine  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: LDP0  
 Center clone name: CH250-26885  
 ----- Summary Statistics  
 Chemistry: Dye-Primer Bodypy: inf of reads  
 Chemistry: Dye-Terminator Big Dye: inf of reads  
 Consensus quality: 3324 bases at least Q40  
 Consensus quality: 3942 bases at least Q30  
 Consensus quality: 4602 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
 \* NOTE: The contigs are based on the application  
 \* of the PGI method using the Human genome (NCBI build 31)  
 \* as the comparative genome.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 1 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

\*  
 1 138875: contig of 138875 bp in length.

**FEATURES**  
 source  
 1. 138875  
 /organism="Macaca mulatta"  
 /mol type="genomic DNA"  
 /db xref="taxon:9544"  
 /clone="CH250-26885"  
 1. 138875  
 /note="assembly name: CH250-26885.1A  
 CONFIDENCE: 0.67"

**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 3.67e-10 Length: 138875  
 Score: 19.00 Matches: 19  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatch: 0  
 Query Match: 13.38% Indels: 0  
 DB: 2 Gaps: 0

US-09-690-825-34 (1-142) x AC142802 (1-138875)

**Cy** 114 AlalysGluThrSerAsnLysLysGluPheGluGluThrAlaLysLysValArg 132  
 |||  
 Db 104968 GCAAGAGAAACCAACATTAAGAGAAATTGAGAAATCTGCAAGAAAGTCGCC 105024

**RESULT 49**  
 AC142803 166874 bp DNA linear HTG 09-APR-2003  
 LOCUS  
 DEFINITION Macaca mulatta clone CH250-26886, \*\*\* SEQUENCING IN PROGRESS \*\*\*.  
 ACCESSION AC142803.1 GI:29567442  
 VERSION  
 HTG: HTGS PHASE2; HTGS PGI.  
 KEYWORDS  
 SOURCE Macaca mulatta (rhesus monkey)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Macaca.

**REFERENCE**  
 1 (bases 1 to 166874)  
 Curoso, M., and Milosavljevic, A.  
 Pooled genomic indexing (PGI): mathematical analysis and experiment  
 design  
 (in) Guigo, R. and Gusfield, D. (Eds.):  
 ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI  
 2002, ROME, ITALY, SEPTEMBER 17-21, 2002. PROCEEDINGS: 10-28;  
 Springer (2002)

**JOURNAL**  
 2 (bases 1 to 166874)  
 Milosavljevic, A., Sodergren, E., Curoso, M., Li, B., Jackson, A.R.,  
 Adams, C., Adio-Odoilo, B., Ali-osman, F.R., Allen, C., Atchcocks, S.L.,  
 Amaratunga, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J.,  
 Benton, J., Biraghe, K., Blankenburg, K., Bonini, D., Bock, U.,  
 Bowle, S., Brileva, M., Brown, S., Brown, M., Bryant, N.P., Bulay, C.,  
 Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,  
 Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,  
 Chen, Z., Chiu, D., Chowdhury, I., Christopoulos, C., Cleveland, C.D.,  
 Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,  
 Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,  
 Dem, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,  
 Dugan-Rocha, S., Durbin, K.J., Egan, A., Ehmhart, C., Edwards, C.C.,  
 Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferriguto, D.,  
 Flagg, N., Ford, J., Foster, P., Frantz, P., Gabris, A., Gao, J.,  
 Garcia, A., Garner, T., Garza, N., Gill, R., Gorell, J.H., Guevara, M.,  
 Garatane, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K.,  
 Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O.,  
 Hodgeson, A., Hughes, M., Holloway, C., Hollins, B., Homsl, F.,  
 Howard, S., Huber, J., Hui, Y., Hui, S., Hume, J., Ioshikhes, I., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvan, J., Kovar, C.,  
 Kravovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, J.C.,  
 Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, M.,  
 Louisedge, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
 Ma, J., Maheshwari, M., Mapa, P., Marondei, I., Martin, R.,  
 Martindale, A., Martinez, E., Massey, E., Mamhney, E., McLeod, M.P.,  
 Medor, M., Mei, G., Merscher, S., Metzker, M., Miller, A., Miner, G.,  
 Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M.,  
 Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N.,  
 Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokweto, S.,  
 Ogun, M., Okumu, G., Oragunye, N., Oyler, R., Pace, A., Payton, B.,  
 Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.,  
 Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M.,  
 Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shim, C.,  
 Shooshitari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A.,  
 Stanley, H., Stone, H., Sutton, A., Svatek, A., Taylor, P., Tamerisa, A.,  
 Tameis, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B.,  
 Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D.,  
 Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R.,  
 Washington, C., Wallington, S., Williams, G., Williamson, A.,  
 Wleczky, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J.,  
 Zorrilla, S., Zuchterlapati, R., Weinstein, G. and Gibbs, R.

**TITLE**  
 Direct Submission

**JOURNAL**  
 Unpublished

**REFERENCE**  
 3 (bases 1 to 166874)





```

* 18080 18813: contig of 734 bp in length
* 18814 19729: contig of 916 bp in length
* 19730 20675: contig of 946 bp in length
* 20676 21721: contig of 1046 bp in length
* 21722 23111: contig of 1390 bp in length
* 23112 25188: contig of 2077 bp in length
* 25189 26588: contig of 1400 bp in length
* 26589 27644: contig of 1056 bp in length
* 27645 29372: contig of 1728 bp in length
* 29373 29495: contig of 123 bp in length
* 29496 30431: contig of 936 bp in length
* 30432 31449: contig of 1018 bp in length
* 31450 32983: contig of 1534 bp in length
* 32984 33162: contig of 179 bp in length
* 33163 33975: contig of 813 bp in length
* 33976 35031: contig of 1056 bp in length
* 35032 35709: contig of 678 bp in length
* 35710 37308: contig of 1599 bp in length
* 37309 38107: contig of 799 bp in length
* 38108 39102: contig of 995 bp in length
* 39103 39591: contig of 489 bp in length
* 39592 41151: contig of 1560 bp in length
* 41152 42842: contig of 1691 bp in length
* 42843 43828: contig of 986 bp in length
* 43829 44009: contig of 181 bp in length
* 44010 44874: contig of 865 bp in length
* 44875 45191: contig of 317 bp in length
* 45192 46147: contig of 956 bp in length
* 46148 47504: contig of 1357 bp in length
* 47505 49091: contig of 1587 bp in length
* 49092 50259: contig of 1168 bp in length
* 50260 51361: contig of 1102 bp in length
* 51362 52631: contig of 1270 bp in length
* 52632 53607: contig of 976 bp in length
* 53608 54820: contig of 1213 bp in length
* 54821 56600: contig of 1780 bp in length
* 56601 58236: contig of 1636 bp in length

```

```

* 58237 59716: gap of unknown length
* 59717 62221: gap of unknown length
* 62222 65085: gap of unknown length
* 65086 66739: gap of unknown length
* 66740 68359: gap of unknown length
* 68360 70014: gap of unknown length
* 70015 71808: gap of unknown length
* 71809 73581: gap of unknown length
* 73582 75574: gap of unknown length
* 75575 78469: gap of unknown length
* 78469 79231: gap of unknown length
* 79232 83240: gap of unknown length
* 83241 86157: gap of unknown length
* 86158 89150: gap of unknown length
* 89151 92671: gap of unknown length
* 92672 95442: gap of unknown length
* 95443 98476: gap of unknown length
* 98477 100815: gap of unknown length
* 100816 103005: gap of unknown length
* 103006 105514: gap of unknown length
* 105515 108704: gap of unknown length
* 108705 111097: gap of unknown length
* 111098 114888: gap of unknown length
* 114889 117423: gap of unknown length

```

## Alignment Scores:

```

Pred. No.: 3,07e-05 Length: 229426
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.56% Indels: 0
DB: 2 Gaps: 0

```

US-09-690-825-34 (1-142) x ACO10532 (1-229426)

Qy 114 AAlaYsGluThrAsnAsnIySlySgluphegiuIurrrAla 128

Db 159273 GCAAGAGAACCAACATATAGAGAAAGAAATTTGAGAAACTGCG 159229

## RESULT 51

AY174765

LOCUS AY174765

DEFINITION Xenopus laevis SIX mRNA, complete cds.

ACCESSION AY174765

VERSION AY174765.1 GI:27762628

## KEYWORDS

SOURCE Xenopus laevis (African clawed frog)

## ORGANISM

Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipodae; Pipidae; Xenopodinae; Xenopus.  
 1 (bases 1 to 772)  
 Song, K.-H., Kim, T.-M., Kim, H.-J., Kim, J.W., Kim, H.-H., Kwon, H.-B., Kim, W.S. and Choi, H.-S.  
 Molecular cloning and characterization of a novel inhibitor of apoptosis protein from *Xenopus laevis*  
 Unpublished  
 2 (bases 1 to 772)  
 Song, K.-H., Kim, T.-M., Kim, H.-J., Kim, J.W., Kim, H.-H., Kwon, H.-B., Kim, W.S. and Choi, H.-S.  
 Direct Submission  
 Submitted (07-NOV-2002) Hormone Research Center, 300 Yongsong-dong Buk-gu, Kwangju 500-757, Korea  
 Location/Qualifiers  
 1..772  
 /organism="Xenopus laevis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8355"  
 6..479  
 /note="survivin in *Xenopus*"  
 /codon\_start=1  
 /product="Six"  
 /protein\_id="AA020085.1"  
 /db\_xref="GI:27762628"  
 /translation="MSTISPLRCDNFPSPDEWRLYNATRLTFTSPFTEDC  
 ACPEPMAEGFVACPTDNPSPVCKFCCKELEGWQPEDPDEHKSPSCFLAL  
 KKKAEEPLSEFLKLDLEHTKIKWQKQKMHIERFQAKANEVGHLEKADADPTQ"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.56e-05 Length: 772  
 Score: 13.00 Matches: 13  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 9.15% Indels: 0  
 Gaps: 0  
 DB: 5  
 US-09-690-825-34 (1-142) x AY174765 (1-772)  
 Cy 31 CysAlaCysThrProGluArgMetAlaGluAlaGlyPhe 43  
 Db 135 TGGCGTTGACCCGAGCGGATGGCAGAGGTGATTT 173  
 RESULT 52 AC141438 207347 bp DNA linear HTG 16-MAR-2003  
 AC141438 Mus musculus chromosome UNK clone RP23-212C9, WORKING DRAFT  
 LOCUS AC141438  
 DEFINITION SEQUENCE, 8 unordered pieces.  
 ACCESSION AC141438.2 GI:28975182  
 VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 207347)  
 McPherson, J.D. and Waterston, R.H.  
 The sequence of Mus musculus clone  
 Unpublished  
 2 (bases 1 to 207347)  
 McPherson, J.D. and Waterston, R.H.  
 Direct Submission  
 Submitted (15-MAR-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 3 (bases 1 to 207347)  
 McPherson, J.D. and Waterston, R.H.  
 Direct Submission  
 Submitted (16-MAR-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 On Mar 16, 2003 this sequence version replaced gi:28973971.  
 COMMENT

----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc/index.shtml>  
 Contact: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
 Project Information  
 Center project name: M\_BA0212C09  
 ----- Summary Statistics -----  
 Sequencing vector: M13; 0%  
 Sequencing vector: plasmid; 100%  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 204828 bases at least Q40  
 Consensus quality: 205544 bases at least Q30  
 Consensus quality: 206016 bases at least Q20  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 8 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 1 1062: contig of 1062 bp in length  
 \* 1063 1162: gap of unknown length  
 \* 1163 5221: contig of 4059 bp in length  
 \* 5222 5321: gap of unknown length  
 \* 5322 16616: contig of 11295 bp in length  
 \* 16617 16716: gap of unknown length  
 \* 16717 31077: contig of 14361 bp in length  
 \* 31078 31177: gap of unknown length  
 \* 31178 59093: contig of 27916 bp in length  
 \* 59094 59193: gap of unknown length  
 \* 59194 74865: contig of 15672 bp in length  
 \* 74866 74966: gap of unknown length  
 \* 74967 128601: contig of 53636 bp in length  
 \* 128602 128701: gap of unknown length  
 \* 128702 207347: contig of 76646 bp in length.  
 Location/Qualifiers  
 source  
 1..207347  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="UNK"  
 /clone="RP23-212C9"  
 1..1062  
 /note="assembly\_name:Contig32"  
 1163..5221  
 /note="assembly\_name:Contig34"  
 5322..16616  
 /note="assembly\_name:Contig36"  
 16717..31077  
 /note="assembly\_name:Contig37"  
 31178..59093  
 /note="assembly\_name:Contig38"  
 59194..74865  
 /note="assembly\_name:Contig39"  
 74966..128601  
 /note="assembly\_name:Contig40"  
 128702..207347  
 /note="assembly\_name:Contig41"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.00628 Length: 207347  
 Score: 13.00 Matches: 13  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 9.15% Indels: 0  
 Gaps: 0  
 DB: 2

US-09-690-825-34 (1-142) x AC141438 (1-207347)

Qy 108 ArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 |||||  
 Db 27135 AGAGCCAGACAAATTCGAAAGAGCAACACACAG 27173

RESULT 53  
 AF442492 483 bp mRNA linear VRT 03-DEC-2002  
 LOCUS Xenopus laevis survivin mRNA, complete cds.  
 DEFINITION AF442492  
 ACCESSION AF442492.1 GI:25990776  
 VERSION  
 KEYWORDS  
 SOURCE Xenopus laevis (African clawed frog)  
 ORGANISM

REFERENCE  
 AUTHORS Murphy, C.R., Sabel, J.L., Sandler, A.D. and Dagle, J.M.  
 TITLE Survivin mRNA is down-regulated during early Xenopus laevis  
 embryogenesis  
 JOURNAL Dev. Dyn. 225 (4), 597-601 (2002)  
 MEDLINE 22341152  
 PUBMED 12454937  
 REFERENCE 2 (bases 1 to 483)  
 AUTHORS Murphy, C.R., Sabel, J.L., Sandler, A.D. and Dagle, J.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-NOV-2001) Pediatrics, University of Iowa, 200 Hawkins  
 Drive, Iowa City, IA 52242, USA  
 LOCATION/Qualifiers

FEATURES  
 source 1..483  
 /organism="Xenopus laevis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8355"  
 /cell\_type="oocyte"  
 /dev\_stage="stage VI"  
 1..483  
 /note="inhibitor of apoptosis"  
 /codon\_start=1  
 /product="survivin"  
 /protein\_id="AA076690.1"  
 /db\_xref="GI:25990776"  
 /translation="MYSANRFEVQAVORLQDFRMYDYEARLATFADPFTENCKCTP  
 ENMKAAGFVHCPTENEPDVACCFCKLEGEPPDDPMENSHSKSVNCGFSLTKCV  
 NDLTMEGFLRLGDRKSFYRKFSFVLYQYVEEWTAATKRLLEFVSQHQSIDLH  
 "

ORIGIN

Alignment Scores:  
 Pred. No.: 0.000243 Length: 483  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 8.45% Indels: 0  
 DB: 5 Gaps: 0

US-09-690-825-34 (1-142) x AF442492 (1-483)

Qy 62 LysGluLeuGluGlyTrpGluProAspAspPro 73  
 |||||  
 Db 211 AAGGAGCTGGAGGCTGGAGCCAGATGATGACCT 246

RESULT 54  
 AY115553 483 bp mRNA linear VRT 17-JUL-2002  
 LOCUS Xenopus laevis survivin mRNA, complete cds.  
 DEFINITION AY115553  
 ACCESSION AY115553.1 GI:21898547  
 VERSION  
 KEYWORDS  
 SOURCE Xenopus laevis (African clawed frog)  
 ORGANISM

REFERENCE  
 AUTHORS Bolton, M.A., Lan, W., Powers, S.E., McClelland, M.L., Kuang, J. and  
 1 (bases 1 to 483)  
 Bolton, M.A., Lan, W., Powers, S.E., McClelland, M.L., Kuang, J. and  
 Stukenberg, P.T.  
 TITLE Aurora B kinase exists in a complex with survivin and INCENP and  
 its kinase activity is stimulated by survivin binding and  
 phosphorylation  
 JOURNAL Mol. Biol. Cell (2002) In press  
 REFERENCE 2 (bases 1 to 483)  
 AUTHORS Bolton, M.A., Lan, W., Powers, S.E., McClelland, M.L., Kuang, J. and  
 Stukenberg, P.T.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-MAY-2002) Biochemistry and Molecular Genetics,  
 University of Virginia Medical School, Health Systems Box 800733,  
 Charlottesville, VA 22908, USA  
 LOCATION/Qualifiers

FEATURES  
 source 1..483  
 /organism="Xenopus laevis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8355"  
 /dev\_stage="11.5-14"  
 1..483  
 /note="aurora B kinase complex subunit; Birt, physically  
 associated with INCENP and aurora B kinase; mitotic  
 regulator of the inner centromere"  
 /codon\_start=1  
 /product="survivin"  
 /protein\_id="AA076714.1"  
 /db\_xref="GI:21898548"  
 /translation="MYSANRFEVQAVORLQDFRMYDYEARLATFADPFTENCKCTP  
 ENMKAAGFVHCPTENEPDVACCFCKLEGEPPDDPMENSHSKSVNCGFSLTKCV  
 NDLTMEGFLRLGDRKSFYRKFSFVLYQYVEEWTAATKRLLEFVSQHQSIDLH  
 "

ORIGIN

Alignment Scores:  
 Pred. No.: 0.000243 Length: 483  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 8.45% Indels: 0  
 DB: 5 Gaps: 0

US-09-690-825-34 (1-142) x AY115553 (1-483)

Qy 62 LysGluLeuGluGlyTrpGluProAspAspPro 73  
 |||||  
 Db 211 AAGGAGCTGGAGGCTGGAGCCAGATGATGACCT 246

RESULT 55  
 AF322051 503 bp mRNA linear VRT 23-AUG-2002  
 LOCUS Gallus gallus survivin mRNA, complete cds.  
 DEFINITION AF322051  
 ACCESSION AF322051.1 GI:11992272  
 VERSION  
 KEYWORDS  
 SOURCE Gallus gallus (chicken)  
 ORGANISM

REFERENCE  
 AUTHORS Johnson, A.L., Langer, J.S. and Brigham, J.T.  
 TITLE Survivin as a Cell Cycle-Related and Antiapoptotic Protein in  
 Granulosa Cells  
 JOURNAL Endocrinology 143 (9), 3405-3413 (2002)  
 PUBMED 12193553  
 REFERENCE 2 (bases 1 to 503)  
 AUTHORS Langer, J.S., Johnson, A.L. and Brigham, J.T.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-NOV-2000) Department of Biology, University of Notre

FEATURES  
Source Dame, P.O. Box 369, Notre Dame, IN 46556, USA  
Location/Qualifiers  
1.503  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
9.437  
/note="inhibitor of apoptosis survivin; BIR domain containing protein"  
/codon\_start=1  
/product="survivin"  
/protein\_id="AAG42494.1"  
/db\_xref="GI:11992273"  
/translation="MAAYAEMLPKEMLVLYVSTRATFRNWPFTGCACTPERMAAG FVHCPSNSPDVACCFCKLEGEWPDDEDDLEHKKHSAGCAFALQKDPNLTVOE FLTKDKRTKNVTKKAISCKETDIEDVAKGVRAIENMGV"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.000253 Length: 503  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.45% Indels: 0  
DB: 5 Gaps: 0

US-09-690-825-34 (1-142) x AF322051 (1-503)

QY 62 lysglulengluglyttrgltuproaspaspapppro 73  
DB 198 AAGGAGCTGGAGGCTGGAGCCGACGACGACCCG 233

RESULT 56 AF377323 555 bp mRNA linear VRT 23-AUG-2002  
AF377323  
LOCUS Gallus gallus survivin delta mRNA, complete cds, alternatively spliced.  
DEFINITION  
ACCESSION AF377323.1 GI:14194233  
VERSION AF377323  
KEYWORDS  
SOURCE Gallus gallus (chicken)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 555)  
AUTHORS Johnson,A.L., Langer,J.S. and Bridgham,J.T.  
TITLES Survivin as a Cell Cycle-Related and Antiapoptotic Protein in Granulosa Cells  
JOURNAL Endocrinology 143 (9), 3405-3413 (2002)  
12193553  
2 (bases 1 to 555)  
REFERENCE Langer,J.S., Bridgham,J.T. and Johnson,A.L.  
TITLES Direct Submission  
JOURNAL Submitted (07-MAY-2001) Biology, University of Notre Dame, P.O. Box 369, Notre Dame, IN 46556, USA  
FEATURES  
Source location/Qualifiers  
1.555  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
1.408  
/note="splice variant containing insertion beyond BIR domain which changes C-terminal predicted survivin"

CDS  
/codon\_start=1  
/product="survivin delta"  
/protein\_id="AAK56308.1"  
/db\_xref="GI:14194234"  
/translation="MAAYAEMLPKEMLVLYVSTRATFRNWPFTGCACTPERMAAG FVHCPSNSPDVACCFCKLEGEWPDDEDDLEHKKHSAGCAFALQKDPNLTVOE FLTKDKRTKNVITRPRGAGCAPLPLALA"

ORIGIN

Alignment Scores:  
Pred. No.: 0.000278 Length: 555  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.45% Indels: 0  
DB: 5 Gaps: 0

US-09-690-825-34 (1-142) x AF377323 (1-555)

QY 62 lysglulengluglyttrgltuproaspaspapppro 73  
DB 190 AAGGAGCTGGAGGCTGGAGCCGACGACGACCCG 225

RESULT 57 AY100639 773 bp mRNA linear VRT 06-DEC-2002  
AY100639  
LOCUS Xenopus laevis survivin/XlaR1 mRNA, complete cds.  
DEFINITION  
ACCESSION AY100639.1 GI:22651694  
VERSION AY100639  
KEYWORDS  
SOURCE Xenopus laevis (African clawed frog)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Xenopus.  
REFERENCE 1 (bases 1 to 773)  
AUTHORS Losada,A., Hirano,M. and Hirano,T.  
TITLES Cohesin release is required for sister chromatid resolution, but not for condensin-mediated compaction, at the onset of mitosis  
JOURNAL Genes Dev. 16 (23), 3004-3016 (2002)  
22352081  
MEDLINE  
PUBMED 12464631  
REFERENCE 2 (bases 1 to 773)  
AUTHORS Hirano,M. and Hirano,T.  
TITLES Direct Submission  
JOURNAL Submitted (30-APR-2002) Cold Spring Harbor Laboratory, One Bungtown Road, Cold Spring Harbor, NY 11724, USA  
location/Qualifiers

FEATURES  
Source location/Qualifiers  
1.773  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
33.515  
/note="regulatory subunit of kinase complex"

CDS  
/codon\_start=1  
/product="survivin/XlaR1"  
/protein\_id="AAW44085.1"  
/db\_xref="GI:22651695"  
/translation="MYSAKNRFVCAVQRLQDFKNMYDYDARLATPFADEFTENCKCTP ESMARAGFVHCPTENSPDVACCFCKLEGEWPDDEDDLEHKKHSAGCAFALQKDPNLTVOE NDLTMEGRLEGRLEKIKSFYKRFSTVLAQYVEEENTATPKRLLEYFSNQHSIDLDD "

ORIGIN

Alignment Scores:  
Pred. No.: 0.000386 Length: 773  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.45% Indels: 0  
DB: 5 Gaps: 0

US-09-690-825-34 (1-142) x AY100639 (1-773)

QY 62 lysglulengluglyttrgltuproaspaspapppro 73  
DB 243 AAGGAGCTGGAGGCTGGAGCCGACGATGATGACCT 278

RESULT 58 AL831742 230800 bp DNA linear ROD 20-AUG-2002  
AL831742/c  
LOCUS Mouse DNA sequence from clone RP23-11F9 on chromosome 2, complete  
DEFINITION

sequence.  
 ACCESSION AL831742  
 VERSION AL831742.6 GI:22449823  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 230800)  
 REFERENCE Tracey, A.  
 Direct Submission  
 Submitted (20-AUG-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail emulites: humquerry@sanger.ac.uk  
 On Aug 22, 2002 this sequence version replaced gi:22449823.  
 JOURNAL Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: humquerry@sanger.ac.uk

## COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPeP; Information on the WormPeP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) R23-11F9 is from the R23-23 Mouse PAC Library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBRAC3.6.

FEATURES  
 source  
 1..230800  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="2"  
 /clone="RP23-11F9"  
 /clone\_1ib="RP23-11F9"  
 /clone\_2ib="RP23-11F9"

## ORIGIN

Alignment Scores:  
 Pred. No.: 23.7 Length: 230800  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 7.04% Indels: 0  
 DB: 10 Gaps: 0

US-09-690-825-34 (1-142) x AL831742 (1-230800)

Qy 132 ArgArgAlaIleGluGlnLeuAlaIleMet 141

DB 57705 AGGAGAGCGATTGACAGCTGCGACGACATG 57676

## RESULT 59

AC109963/c

LOCUS AC109963 266342 bp DNA linear HTG 18-SEP-2002  
 DEFINITION Rattus norvegicus clone CH230-139E1, \*\*\* SEQUENCING IN PROGRESS  
 ACCESSION AC109963

AC109963

## VERSION

AC109963.4 GI:23101152

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

## REFERENCE

1 (bases 1 to 266342)  
 AUTHORS Muzny, D., Maric, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angilano, D., Anyalebech, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Behnke, F., Biswal, K., Blatt, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunney, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy, Carroll, L., De Andrade, C., Dedrich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabiella, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hognes, M., Hollins, B., Howell, S., Huijck, S., Hume, J., Idledit, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyvet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kows, C., Kraft, C.L., Labow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, M., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzini, L., Louisedge, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Maindarte, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Manthey, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, D., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okunolu, G., Olampunagbon, A., Pal, S., Parke, K., Pascherak, S., Paul, H., Perez, A., Perez, A., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinhilber, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wiczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

## JOURNAL

Unpublished

2 (bases 1 to 266342)

## AUTHORS

Worley, K.C.

## TITLE

Submitted (09-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

## REFERENCE

Submitted (18-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

## JOURNAL

Unpublished

## COMMENT

The sequence in this assembly is a combination of BAC based reads



## ORGANISM

Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.

REFERENCE  
AUTHORS  
TITLE  
1 (bases 1 to 316)  
Johnson, A.L., Langer, J.S. and Bridgman, J.T.  
Survivin as a Cell Cycle-Related and Antipoptotic Protein in  
Granulosa Cells

JOURNAL  
PUBMED  
12193553  
2 (bases 1 to 316)  
Langer, J.S., Bridgman, J.T. and Johnson, A.L.

REFERENCE  
AUTHORS  
TITLE  
Submitted (07-MAY-2001) Biology, University of Notre Dame, P.O. Box  
369, Notre Dame, IN 46556, USA

## FEATURES

source  
Location/Qualifiers

## CDS

1..180  
/note="splice variant containing deletion within BIR  
domain causing early truncation of predicted survivin"  
/codon\_start=1  
/product="survivin short"  
/protein\_id="AAK56307.1"  
/db\_xref="GI:14194232"  
/translation="MAAYAEMLPKEMLVYVSTRATFRMPFTEGCACTPERGCTOK  
ALRGARFCSSRSRL"

## ORIGIN

## Alignment Scores:

Pred. No.: 0.543 Length: 316  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.34% Indels: 0  
DB: 5 Gaps: 0

US-09-690-825-34 (1-142) x AF377322 (1-316)

QY 29 GIUGIYCYGALACySThPrGGLuArg 37

Db 91 GAGGCTGCGCCTGCACGCCCGAGCGG 117

## RESULT 63

AF377324

LOCUS 445 bp mRNA linear VRT 23-NOV-2002  
DEFINITION Gallus gallus survivin gamma mRNA, complete cds, alternatively  
spliced.

ACCESSION AF377324

VERSION AF377324.1 GI:14194235

KEYWORDS Gallus gallus (chicken)

## SOURCE

Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.

REFERENCE  
AUTHORS  
TITLE  
1 (bases 1 to 445)  
Johnson, A.L., Langer, J.S. and Bridgman, J.T.  
Survivin as a Cell Cycle-Related and Antipoptotic Protein in  
Granulosa Cells

JOURNAL  
PUBMED  
12193553  
2 (bases 1 to 445)  
Langer, J.S., Bridgman, J.T. and Johnson, A.L.

REFERENCE  
AUTHORS  
TITLE  
Submitted (07-MAY-2001) Biology, University of Notre Dame, P.O. Box  
369, Notre Dame, IN 46556, USA

## FEATURES

source  
Location/Qualifiers  
1..445  
/organism="Gallus gallus"  
/mol\_type="mRNA"

## CDS

/db\_xref="taxon:9031"  
1..180  
/note="splice variant containing deletion within BIR  
domain and insertion beyond BIR domain; deletion results  
in early truncation of predicted survivin"  
/codon\_start=1  
/product="survivin gamma"  
/protein\_id="AAK56309.1"  
/db\_xref="GI:14194236"  
/translation="MAAYAEMLPKEMLVYVSTRATFRMPFTEGCACTPERGCTOK  
ALRGARFCSSRSRL"

## ORIGIN

## Alignment Scores:

Pred. No.: 0.761 Length: 445  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.34% Indels: 0  
DB: 5 Gaps: 0

US-09-690-825-34 (1-142) x AF377324 (1-445)

QY 29 GIUGIYCYGALACySThPrGGLuArg 37

Db 91 GAGGCTGCGCCTGCACGCCCGAGCGG 117

## RESULT 64

AY057058

LOCUS 489 bp mRNA linear VRT 29-OCT-2001  
DEFINITION Danio rerio survivin 2 (b1r5b) mRNA, complete cds.

ACCESSION AY057058

VERSION AY057058.1 GI:16508147

## KEYWORDS

Danio rerio (zebrafish)

## SOURCE

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.

## REFERENCE

1 (bases 1 to 489)  
Inohara, N. and Nunez, G.

Genes with homology to mammalian apoptosis regulators identified in  
zebrafish

JOURNAL

Cell Death Differ. 7 (5), 509-510 (2000)

PUBMED

10917738

## TITLE

2 (bases 1 to 489)  
Inohara, N. and Nunez, G.

Direct Submission

Submitted (24-SEP-2001) Department of Pathology, University of  
Michigan Medical Center, 1500 E. Medical Center Dr., Ann Arbor, MI  
48109, USA

## FEATURES

source  
Location/Qualifiers  
1..489  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
1..489  
/gene="b1r5b"  
27..413  
/gene="b1r5b"  
/codon\_start=1  
/product="survivin 2"  
/protein\_id="AA18251.1"  
/db\_xref="GI:16508148"  
/translation="MYSYKRIQTFSEWPFEDQCCTPELMAKGFVHCPSENEBDA  
CCFYLRLEGWEPDNDPMSHAKSPCAFLHMSKTPDELTALEYFHLDEORLIYI  
KMGGRKAYPDEVATSKQLRALI"

## ORIGIN

Alignment Scores:  
Pred. No.: 0.835 Length: 489  
Score: 9.00 Matches: 9



Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.34% Indels: 0  
 DB: 5 Gaps: 0

US-09-690-825-34 (1-142) X AY057058 (1-489)

QY 63 GluLeuGluGlyTyrGluProAspAsp 71  
 Db 180 GAACTGGAGGCTGGGAAACGATGAC 206

RESULT 65  
 AC107911/c 88863 bp DNA linear PRI 25-FEB-2003  
 LOCUS Homo sapiens chromosome 17, clone CTD-2507J6, complete sequence.  
 AC107911  
 AC107911.21 GI:28557953  
 HGK.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 88863)  
 Birren,B., Nusbaum,C. and Lander,E.  
 Homo sapiens chromosome 17, clone CTD-2507J6  
 Unpublished  
 2 (bases 1 to 88863)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
 Brown,A., Camarata,J., Campopiano,A., Chang,U., Chazaro,B.,  
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,  
 Cooke,P., DeRellano,K., Dewar,X., Diaz,J.S., Dodge,S., Faro,S.,  
 Ferreira,P., Fitzhugh,W., Dewar,X., Galagan,J., Gardyna,S.,  
 Glude,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Lander,E., Levine,R.,  
 Lander,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,  
 MacDonald,P., Major,J., Margus,N., Matthews,C., McCarthy,M.,  
 Mcwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,  
 Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,  
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,  
 Peterson,K., Phunhngang,P., Pierre,N., Pollara,V., Raymond,C.,  
 Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,U.,  
 Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S.,  
 Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Tophan,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 88863)  
 Birren,B., Nusbaum,C., Lander,E., Abouellell,A., Allen,N.,  
 Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
 Boguslavsky,L., Boukhgalter,B., Camarata,U., Chang,U., Choepel,Y.,  
 Collymore,A., Cook,A., Cooke,P., Cotum,B., DeRellano,K.,  
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
 Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Lander,E., Levine,R.,  
 Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,  
 MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
 Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,  
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Oliver,J., Peterson,K., Phunhngang,P., Pierre,N.,  
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,U., Schauer,S., Schnupack,R., Seaman,S., Severy,P.,  
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Tophan,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.

TITLE  
 JOURNAL  
 RESEARCH  
 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 88863)  
 Birren,B., Nusbaum,C., Lander,E., Abouellell,A., Allen,N.,  
 Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
 Boguslavsky,L., Boukhgalter,B., Camarata,U., Chang,U., Choepel,Y.,  
 Collymore,A., Cook,A., Cooke,P., Cotum,B., DeRellano,K.,  
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
 Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Lander,E., Levine,R.,  
 Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,  
 MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
 Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,  
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Oliver,J., Peterson,K., Phunhngang,P., Pierre,N.,  
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,U., Schauer,S., Schnupack,R., Seaman,S., Severy,P.,  
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Tophan,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.

COMMENT  
 Direct Submission  
 Submitted (25-FEB-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Feb 25, 2003 this sequence version replaced gi:27819569.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE  
 JOURNAL  
 RESEARCH  
 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 88863)  
 Birren,B., Nusbaum,C., Lander,E., Abouellell,A., Allen,N.,  
 Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
 Boguslavsky,L., Boukhgalter,B., Camarata,U., Chang,U., Choepel,Y.,  
 Collymore,A., Cook,A., Cooke,P., Cotum,B., DeRellano,K.,  
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
 Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Lander,E., Levine,R.,  
 Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,  
 MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
 Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,  
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Oliver,J., Peterson,K., Phunhngang,P., Pierre,N.,  
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,U., Schauer,S., Schnupack,R., Seaman,S., Severy,P.,  
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Tophan,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.

FEATURES  
 source  
 1..88863  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17"  
 /clone="CTD-2507J6"  
 /clone\_1b="CTD2 Human BAC"  
 1..107  
 /rpt\_family="AluSg/x"  
 112..404  
 /rpt\_family="AluSx"  
 complement(494..627)  
 /rpt\_family="L2"  
 1264..1559  
 /rpt\_family="AluSx"  
 1566..1870  
 /rpt\_family="AluSx"  
 1955..2080  
 /rpt\_family="MIR"  
 2306..2361  
 /rpt\_family="MIR"  
 complement(5981..6144)  
 /rpt\_family="FAM"  
 6145..6281  
 /rpt\_family="MIR"  
 complement(6343..6663)  
 /rpt\_family="AluSx"  
 complement(6824..6982)  
 /rpt\_family="MIR3"  
 complement(7984..8053)  
 /rpt\_family="MIR"

```

repeat_region complement(8055, .6353)
/rpt_family="AlusX"
repeat_region 8424, .8973
/rpt_family="MLT2B1"
repeat_region 11586, .11608
/rpt_family="GC_rich"
repeat_region 12914, .13033
/rpt_family="L2"
repeat_region 14640, .14950
/rpt_family="MLT1U"
repeat_region 15224, .15333
/rpt_family="L1PA8"
repeat_region 15334, .15647
/rpt_family="AluY"
repeat_region 15648, .15660
/rpt_family="L1PA8"
repeat_region 15753, .16062
/rpt_family="AlusG"
repeat_region 16112, .16224
/rpt_family="L1MC4"
repeat_region 16674, .17106
/rpt_family="MLT1C"
repeat_region 17711, .17835
/rpt_family="F1AM_C"
repeat_region 17840, .17880
/rpt_family="CA1"
repeat_region complement(18017, .18316)
/rpt_family="AlusG"
repeat_region 18318, .18600
/rpt_family="L1MC4"
repeat_region 18601, .18859
/rpt_family="AlusB"
repeat_region 18900, .19245
/rpt_family="L1MC4a"
repeat_region 19640, .19773
/rpt_family="AlusG/x"
repeat_region 19774, .20064
/rpt_family="AlusX"
repeat_region 20065, .20095
/rpt_family="AT_rich"
repeat_region 20389, .20656
/rpt_family="MIR"
repeat_region complement(21003, .21068)
/rpt_family="AlusB/g"
repeat_region 21235, .21532
/rpt_family="AlusX"
repeat_region complement(21533, .21629)
/rpt_family="L1"
repeat_region complement(21631, .21843)
/rpt_family="T19ger1"
repeat_region complement(21844, .22197)
/rpt_family="AlusX"
repeat_region complement(22198, .22707)
/rpt_family="T19ger1"
repeat_region complement(22708, .23007)
/rpt_family="AlusG"
repeat_region complement(23008, .23361)
/rpt_family="T19ger1"
repeat_region complement(23570, .23869)
/rpt_family="AlusX"
repeat_region 23895, .24115
/rpt_family="L1MB3"
repeat_region complement(24132, .24443)
/rpt_family="AluT0"
repeat_region complement(24481, .24905)
/rpt_family="L1R61"
repeat_region complement(24906, .25217)

```

Alignment Scores:

Pred. No.:	139	Length:	88883
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

```

Query Match: 6.34% Indels: 0
DB: 9 Gaps: 0
US-09-690-825-34 (1-142) x AC107911 (1-88883)
Cy 130 LysValArgArgAlaIleGluGluLeu 138
Db 60477 AAGGTAGAGAGACTATTGACGAGCTC 60451

RESULT 66
AC016292 104435 bp DNA linear PRI 19-MAR-2003
LOCUS Homo sapiens chromosome 17, clone RP11-22G12, complete sequence.
DEFINITION
AC016292
VERSION
AC016292.26 GI:29124200
KEYWORDS
HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 104435)
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Homo sapiens chromosome 17, clone RP11-22G12
Unpublished
2 (bases 1 to 104435)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeRubeis,A., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galligan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howard,J.C., Johnson,R., Jones,C., Kamm,L., Karatas,A., Klein,J.,
Lehoczky,J., Lien,C., Locke,K., MacDonald,P., Marquis,N.,
McGowan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 104435)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeRubeis,A.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galligan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamm,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Menes,L., Mihova,T., Monaco,T., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Riese,C., Rogov,P.,
Roman,V., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (31-JAN-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 104435)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeRubeis,A.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galligan,J., Gardyna,S.,

```

TITLE  
JOURNAL

COMMENT  
All repeats were identified using RepeatMasker:  
Smit, A.P.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L4116  
Center clone name: 22\_G\_12

Only the first 104.4 Kilobases of this clone are being submitted.  
The remainder overlaps accession number AC015884 [WIGR project  
L731].

## FEATURES

Source Location/Qualifiers  
1. 104435  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="RP11-22G12"  
/clone\_1lb="RPC1-11 Human Male BAC"  
repeat\_region complement(2. .405)  
repeat\_region /rpt\_family="MLT1J"  
repeat\_region complement(1124. .1507)  
repeat\_region /rpt\_family="AluY"  
repeat\_region complement(1565. .1716)  
repeat\_region /rpt\_family="MIR"  
repeat\_region complement(1745. .1925)  
repeat\_region /rpt\_family="AluSg/x"  
1863  
/note="probably C; possibly T"  
repeat\_region 2174. 2453  
repeat\_region /rpt\_family="AluXx"  
2455. .2481  
repeat\_region /rpt\_family="(CAA)n"  
repeat\_region complement(2901. .3056)  
repeat\_region /rpt\_family="MIR"  
3062. .3120  
repeat\_region /rpt\_family="MIR3"  
3186. .3397  
repeat\_region /rpt\_family="MIR"  
3435. .3522  
repeat\_region /rpt\_family="AluSg/x"  
3523. .3830  
repeat\_region /rpt\_family="AluXx"  
3831. 4079  
repeat\_region /rpt\_family="MIR"  
4912. .4980  
repeat\_region /rpt\_family="Charlie8"  
5105. .5316  
repeat\_region /rpt\_family="MIR"

repeat_region	complement(5324. .5995)
repeat_region	/rpt_family="SVA"
repeat_region	5936. .6246
repeat_region	/rpt_family="AluSx"
repeat_region	complement(6273. .6579)
repeat_region	/rpt_family="AluSx"
repeat_region	complement(6705. .6997)
repeat_region	/rpt_family="AluSg"
repeat_region	7333. .7628
repeat_region	/rpt_family="AluSg"
repeat_region	7827. .8119
repeat_region	/rpt_family="AluSg"
repeat_region	complement(8600. .8909)
repeat_region	/rpt_family="AluSx"
repeat_region	complement(8910. .9043)
repeat_region	/rpt_family="AluSg/x"
repeat_region	complement(9050. .9370)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(9421. .9718)
repeat_region	/rpt_family="AluSx"
repeat_region	9782. .9976
repeat_region	/rpt_family="MIR"
repeat_region	complement(10186. .10519)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(11168. .11288)
repeat_region	/rpt_family="AluSg/x"
repeat_region	11980. .12243
repeat_region	/rpt_family="AluSx"
repeat_region	complement(12613. .12910)
repeat_region	/rpt_family="AluY"
repeat_region	13447. .13757
repeat_region	/rpt_family="AluSx"
repeat_region	complement(13846. .14149)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(14173. .14226)
repeat_region	/rpt_family="L1ME"
repeat_region	complement(14256. .14414)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(14415. .14714)
repeat_region	/rpt_family="AluSx"
repeat_region	complement(14715. .14750)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(14818. .14873)
repeat_region	/rpt_family="AluSg"
repeat_region	14981. .15127
repeat_region	/rpt_family="PRAM"
repeat_region	complement(15128. .15200)
repeat_region	/rpt_family="L1M4"
repeat_region	complement(15302. .15360)
repeat_region	/rpt_family="AluYb"
repeat_region	complement(15384. .15550)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(15551. .15854)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(15855. .15988)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(16173. .16287)
repeat_region	/rpt_family="MIR"
repeat_region	16318. .16625
repeat_region	/rpt_family="AluSx"
repeat_region	complement(16644. .16712)
repeat_region	/rpt_family="MER7A"
repeat_region	complement(16782. .17002)
repeat_region	/rpt_family="MER7A"
repeat_region	complement(17043. .17354)

Alignment Scores:  
Pred. No.: 163  
Score: 9.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 6.34%  
DB: 9  
Gaps: 0

Length: 104435  
Matches: 9  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-690-825-34 (1-142) x AC016292 (1-104435)

QY 130 LysValArgArgAlaIleGluGlnLeu 138  
 DB 47584 AAGGTAGAGAGAGCTATTGACGAGCTC 47610

RESULT 67  
 AC145222

LOCUS AC145222 136470 bp DNA linear HTG 04-DEC-2003  
 DEFINITION Medicago truncatula clone mth2-29a15, WORKING DRAFT SEQUENCE, 7  
 unorderded pieces.  
 AC145222  
 AC145222.7 GI:38678560  
 HTG; HTGS PHASE1; HTGS DRAFT.  
 SOURCE Medicago truncatula (barrel medic)  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 136470)  
 AUTHORS Shall,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.  
 TITLE Medicago truncatula BAC Clone mth2-29a15  
 JOURNAL Unpublished  
 2 (bases 1 to 136470)  
 AUTHORS Shall,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-JUN-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA  
 3 (bases 1 to 136470)  
 AUTHORS Shall,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-DEC-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA  
 COMMENT ON Dec 4, 2003 this sequence version replaced gi:38502390.

----- Genome Center  
 Center: Department Of Chemistry And Biochemistry  
 The University Of Oklahoma  
 Center code:UOKNOR

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 7 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 3444: contig of 3444 bp in length  
 3445 3544: gap of unknown length  
 3545 10199: contig of 6655 bp in length  
 10200 10299: gap of unknown length  
 10300 22748: contig of 12449 bp in length  
 22749 22848: gap of unknown length  
 22849 39691: contig of 16843 bp in length  
 39692 39792: gap of unknown length  
 39793 65373: contig of 25482 bp in length  
 65374 92086: contig of 26712 bp in length  
 92086 92185: gap of unknown length  
 92186 136470: contig of 44285 bp in length.

Location/Qualifiers  
 1. 136470  
 /organism="Medicago truncatula"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3880"  
 /clone="mth2-29a15"

ORIGIN /clone\_lib="Medicago truncatula BAC library H2"

Alignment Scores:  
 Pred. No.: 213 Length: 136470  
 Score: 9.00 Matches: 9  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.34% Indels: 0  
 DB: 2 Gaps: 0

US-09-690-825-34 (1-142) x AC145222 (1-136470)

QY 115 LysGluThrAspAsnLysLysLysGlu 123  
 DB 130497 AAGAGACTAATTAACAGAGAAAGAA 130523

RESULT 68  
 AC146981/c

LOCUS AC146981 146515 bp DNA linear HTG 30-OCT-2003  
 DEFINITION Homo sapiens chromosome 17 clone RP13-856H21 map 17, WORKING DRAFT  
 SEQUENCE, 22 unorderded pieces.  
 AC146981  
 AC146981.1 GI:38044233  
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 146515)  
 AUTHORS Birren,B., Nusbaum,C. and Lander,E.  
 TITLE Homo sapiens chromosome 17, clone RP13-856H21  
 JOURNAL Unpublished  
 2 (bases 1 to 146515)  
 AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelell,A., Allen,N., Anderson,N., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,Y., Choepel,Y., Collamore,A., Cook,A., Cooke,P., Cortum,B., Deatellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fayo,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardina,S., Graham,L., Grand-Pierre,N., Hafer,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karac,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Lin,G., Lui,X., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,Y., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vaselliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
 Submitted (30-OCT-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WPIR  
 Web site: http://www-seg.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L30056  
 Center clone name: 856.H.21

----- Summary Statistics  
 Sequencing vector: M13; M77815. 50% of reads  
 Sequencing: plasmid, n/a, 50% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731

Consensus quality: 138962 bases at least Q40  
 Consensus quality: 141529 bases at least Q30  
 Consensus quality: 142778 bases at least Q20  
 Insert size: 197000; agarose-fp  
 Insert size: 144415; sum-of-contigs  
 Quality coverage: 8.7 in Q20 bases; agarose-fp  
 Quality coverage: 11.8 in Q20.  
 NOTE: This is a 'working draft' sequence. It currently  
 consists of 22 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 16866: contig of 16866 bp in length  
 16867 16966: gap of 100 bp  
 16967 17387: contig of 421 bp in length  
 17388 17487: gap of 100 bp  
 17488 17552: contig of 65 bp in length  
 17553 17652: gap of 100 bp  
 17653 18451: contig of 799 bp in length  
 18452 18551: gap of 100 bp  
 18552 19368: contig of 1017 bp in length  
 19369 19669: gap of 100 bp  
 19670 20497: contig of 829 bp in length  
 20498 20598: gap of 100 bp  
 20599 21462: contig of 865 bp in length  
 21463 21562: gap of 100 bp  
 21563 22396: contig of 834 bp in length  
 22397 22496: gap of 100 bp  
 22497 22855: contig of 359 bp in length  
 22856 22955: gap of 100 bp  
 22956 24036: contig of 1081 bp in length  
 24037 24136: gap of 100 bp  
 24137 26011: contig of 1875 bp in length  
 26012 26111: gap of 100 bp  
 26112 28326: contig of 2215 bp in length  
 28327 28426: gap of 100 bp  
 28427 30147: contig of 1721 bp in length  
 30148 30247: gap of 100 bp  
 30248 31844: contig of 1597 bp in length  
 31845 31944: gap of 100 bp  
 31945 34909: contig of 2965 bp in length  
 34910 35009: gap of 100 bp  
 35010 36835: contig of 1826 bp in length  
 36836 36935: gap of 100 bp  
 36936 48373: contig of 11438 bp in length  
 48374 48473: gap of 100 bp  
 48474 61952: contig of 13479 bp in length  
 61953 62052: gap of 100 bp  
 62053 76344: contig of 14292 bp in length  
 76345 76444: gap of 100 bp  
 76445 91535: contig of 15091 bp in length  
 91536 91635: gap of 100 bp  
 91636 109845: contig of 18210 bp in length  
 109846 109945: gap of 100 bp  
 109946 146515: contig of 36570 bp in length.

## FEATURES

## Source

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17"  
 /clone="RP13-856H21"  
 /clone\_id="RPCT-13 Human Female BAC"  
 1..16866  
 /note="assembly\_fragment"  
 clone\_end:56  
 vector\_side:left"  
 16967..17387  
 /note="assembly\_fragment"

## misc\_feature

## misc\_feature

misc\_feature 17488..17552  
 /note="assembly\_fragment"  
 misc\_feature 17653..18451  
 /note="assembly\_fragment"  
 misc\_feature 18552..19368  
 /note="assembly\_fragment"  
 misc\_feature 19669..20497  
 /note="assembly\_fragment"  
 misc\_feature 20598..21462  
 /note="assembly\_fragment"  
 misc\_feature 21563..22396  
 /note="assembly\_fragment"  
 misc\_feature 22497..22855  
 /note="assembly\_fragment"  
 misc\_feature 22956..24036  
 /note="assembly\_fragment"  
 misc\_feature 24137..26011  
 /note="assembly\_fragment"  
 misc\_feature 26112..28326  
 /note="assembly\_fragment"  
 misc\_feature 28427..30147  
 /note="assembly\_fragment"  
 misc\_feature 30248..31844  
 /note="assembly\_fragment"  
 misc\_feature 31945..34909  
 /note="assembly\_fragment"  
 misc\_feature 35010..36835  
 /note="assembly\_fragment"  
 misc\_feature 36936..48373  
 /note="assembly\_fragment"  
 misc\_feature 48474..61952  
 /note="assembly\_fragment"  
 misc\_feature 62053..76344  
 /note="assembly\_fragment"  
 misc\_feature 76445..91535  
 /note="assembly\_fragment"  
 misc\_feature 91636..109845  
 /note="assembly\_fragment"  
 misc\_feature 109946..146515  
 /note="assembly\_fragment"

## ORIGIN

## Alignment Scores:

Pred. No.: 228 Length: 146515  
 Score: 9.00 Matches: 9  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.34% Indels: 0  
 DB: 2 Gaps: 0

US-09-690-825-34 (1-142) x AC146981 (1-146515)

QY 130 lvsValARGATGAlleGluGlnleu 138

DB 110664 AAGGTGAGGAGCTATTGACGAGCTC 110638

## RESULT 69

## LOCUS

## DEFINITION

BX649590 180849 bp DNA linear HTG 31-OCT-2003  
 BX649590  
 Dario reio clone CH211-14316, WORKING DRAFT SEQUENCE, 3 unordered  
 pieces.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

BX649590.3 GI:37606081  
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.  
 Dario reio (zebrafish)  
 Dario reio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 180849)  
 McLaren,S.  
 Direct Submission  
 Submitted (30-OCT-2003) Wellcome Trust Sanger Institute, Hinxton,

Cambridgehire, CB10 1SA, UK. E-mail enquiries:  
zfish-hel@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Oct 9, 2003 this sequence version replaced g1:37496403.

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: zfish-hel@sanger.ac.uk

----- Project Information  
Center project name: zc14336

----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 180245 bases at least Q40  
Consensus quality: 180416 bases at least Q20  
Consensus quality: 180499 bases at least Q20  
Insert size: 180643; sum-of-contigs  
Insert size: 179152; 5.1% error; agarose-fp  
Quality coverage: 9.17x in Q20 bases; sum-of-contigs Quality  
coverage: 9.40x in Q20 bases; agarose-fp

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 3075: contig of 3075 bp in length  
\* 3076 3175: gap of 100 bp  
\* 3176 93133: contig of 89958 bp in length  
\* 93134 93234: gap of 100 bp  
\* 93234 180849: contig of 87616 bp in length.

## FEATURES

## source

1. 180849  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="CH211-14316"  
/clone\_id="CHOR1-211"  
1. 3075

## misc\_feature

/note="assembly fragment:01583  
fragment chain:1  
clone\_end:177  
vector\_side:left"

## misc\_feature

/note="assembly fragment:00359  
fragment chain:1"

## misc\_feature

/note="assembly fragment:01693  
fragment chain:1  
clone\_end:56  
vector\_side:right"

## ORIGIN

## Alignment Scores:

Pred. No.: 280 Length: 180849  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.34% Indels: 0  
Gaps: 0  
DB: 2

US-09-690-825-34 (1-142) x BX645590 (1-180849)

QY 115 LysGIUTHTAsrAsnLYsLYsGLu 123  
|||||  
|||||

DB 20152 AAAGAAACAAATTAATAAAAAGAAAGA 20178  
|||||  
|||||

RESULT 70  
AC015532/C  
LOCUS AC015532 184592 bp DNA linear HTG 26-MAY-2001

## DEFINITION

Homo sapiens chromosome 18 clone RP11-497J22 map 18, WORKING DRAFT

## SEQUENCE

AC015532.4 GI:14209776  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT



JOURNAL  
 Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 184717)  
 AUTHORS  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
 Boukhgalter, B., Brown, A., Camarata, J., Campolongo, A., Chang, J.,  
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,  
 Cook, A., Cooke, P., Dekrellano, K., Dewar, K., Diaz, J., Dodge, S.,  
 Fero, S., Ferrel, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,  
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,  
 Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
 Maclean, C., Macdonald, P., Major, J., Marcus, N., Matthews, C.,  
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L.,  
 Minova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,  
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strause, N., Subramanian, A., Talama, J., Testfaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G.,  
 Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

TITLE  
 JOURNAL  
 Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 184717)  
 AUTHORS  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
 Boukhgalter, B., Brown, A., Camarata, J., Campolongo, A., Chang, J.,  
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,  
 Cook, A., Cooke, P., Dekrellano, K., Dewar, K., Diaz, J., Dodge, S.,  
 Fero, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D.,  
 Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,  
 Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,  
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,  
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K.,  
 Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N.,  
 Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J.,  
 Menus, L., Minova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
 Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P.,  
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,  
 Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,  
 Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,  
 Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Strause, N., Subramanian, A., Talama, J., Testfaye, S.,  
 Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J.,  
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.,  
 Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

TITLE  
 JOURNAL  
 Submitted (15-JUN-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jun 15, 2002 this sequence version replaced g1:20128507.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: MIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L2462  
 Center clone name: 534\_N\_16  
 Location/Qualifiers  
 1. 184717  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="18"

```

/map="18"
/clone="RP11-534N16"
/clone_id="RP11-534N16"
repeat_region
  /rpt_family="MERSA"
  /rpt_family="MERSA"
  complement(1195..1492)
  /rpt_family="AluXs"
  complement(1847..2078)
  /rpt_family="MLTID"
  complement(2190..2369)
  /rpt_family="LIME"
  /rpt_family="LIME"
  complement(2539..2702)
  /rpt_family="LIME"
  complement(2723..2849)
  /rpt_family="L1"
  2895..2945
  /rpt_family="AT-rich"
  complement(2946..3602)
  /rpt_family="L1"
  complement(3616..4019)
  /rpt_family="L1M4"
  complement(4015..4167)
  /rpt_family="L1ME1"
  complement(4204..4687)
  /rpt_family="L1ME1"
  complement(4702..4824)
  /rpt_family="L1PA16"
  complement(4825..5118)
  /rpt_family="AluSg"
  complement(5119..6451)
  /rpt_family="L1PA16"
  complement(6456..6600)
  /rpt_family="MLT1A1"
  7027..7118
  /rpt_family="MIR"
  7395..7431
  /rpt_family="AT-rich"
  complement(8709..9010)
  /rpt_family="AluSg"
  9191..9502
  /rpt_family="AluY"
  9930..10280
  /rpt_family="L1R16C"
  10792..10816
  /rpt_family="(T)n"
  10893..11009
  /rpt_family="L2"
  11010..11302
  /rpt_family="AluUo"
  11303..11857
  /rpt_family="L2"
  complement(11954..11994)
  /rpt_family="L2"
  12294..12650
  /rpt_family="TBE1C"
  13320..13488
  /rpt_family="GA-rich"
  20170..26197
  /rpt_family="L1PA2"
  complement(26602..26873)
  /rpt_family="L2"
  complement(27283..27679)
  /rpt_family="L2"
  28343..28378
  /rpt_family="AT-rich"
  28695..28804
  /rpt_family="U6"
  30955..31267
  /rpt_family="AluUb"
  31641..31794
  /rpt_family="GA-rich"

```



```

repeat_region      31938. .33242
                    /rpt_family="AluSc"
repeat_region      32854. .32915
                    /rpt_family="J3"
repeat_region      complement(33204. .33309)
                    /rpt_family="MIR"
repeat_region      33359. .33594
                    /rpt_family="J2"
repeat_region      complement(33516. .33841)
                    /rpt_family="AluSc"
repeat_region      34590. .34615
                    /rpt_family="AT_rich"
repeat_region      complement(34952. .35106)
                    /rpt_family="T1MC/D"
repeat_region      35119. .35262
                    /rpt_family="J2"
repeat_region      complement(35384. .35559)
                    /rpt_family="MIR"
repeat_region      36614. .36924
                    /rpt_family="AluSc"
repeat_region      38297. .38326
                    /rpt_family="(TATRA)n"
repeat_region      38615. .38662
                    /rpt_family="(TA)n"

Alignment Scores:
Pred. No.:          286
Score:              9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:        6.34%
DB:                  9
                    Gaps: 0

US-09-690-825-34 (1-142) X AC009717 (1-184717)

QY      96 leuhtlenglygluphelenlyleu 104
Db      161494 TTAACATTAGGTGAATCTTAAACCTT 161520

RESULT 72
AC122238      188476 bp   DNA   linear   ROD 25-NOV-2003
LOCUS        Mus musculus BAC clone RP23-143D4 from chromosome 10, complete
DEFINITION
ACCESSION    AC122238
VERSION      AC122238.4 GI:37361064
KEYWORDS     HTG.
SOURCE       Mus musculus (house mouse)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 188476)
              Swearingen-Shahid,S. and Bielicki,L.
              The sequence of Mus musculus BAC clone RP23-143D4
              Unpublished (2001)
REFERENCE    2 (bases 1 to 188476)
              Wilson,R.
              Sequencing of Mus musculus
              Unpublished (2001)
REFERENCE    3 (bases 1 to 188476)
              McPherson,J.D. and Waterston,R.H.
              Direct Submission
              Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE    4 (bases 1 to 188476)
              McPherson,J.D. and Waterston,R.H.
              Direct Submission
              Submitted (19-MAR-2003) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE    5 (bases 1 to 188476)
              Wilson,R.K.
              Direct Submission
              Submitted (02-OCT-2003) Genome Sequencing Center, 4444 Forest Park

```

```

REFERENCE      Parkway, St. Louis, MO 63108, USA
AUTHORS       6 (bases 1 to 188476)
              Wilson,R.
TITLE         Direct Submission
JOURNAL       Submitted (25-NOV-2003) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
              On Oct 2, 2003 this sequence version replaced gi:29124323.
COMMENT       -----
              Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu
              Contact: submissions@watson.wustl.edu
              -----
              Summary Statistics
              -----
              Center project name: M_BA0143D04
              -----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RP23-23 BAC library has been constructed by Kazutoyo Osegawa
and Minako Tateo in the laboratory of Pieter de Jong
(http://www.chori.org) from female C57BL/6J mouse kidney and/or
brain genomic DNA. The clone and detailed information can be
obtained from Research Genetics, Inc. (http://www.resgen.com) or
Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is
overlapped by AC122349.
Location/Qualifiers
1. 188476
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="10"
/map="10"
/clone="RP23-143D4"
/clone_lib="RP23-23"
repeat_region 1854. .2004
/rpt_family="MaLR"
repeat_region 2049. .2197
/rpt_family="Alu"
repeat_region 2244. .2391
/rpt_family="MaLR"
repeat_region 3040. .3222
/rpt_family="B2"
repeat_region 3236. .3404
/rpt_family="Alu"
repeat_region 3452. .3615
/rpt_family="B4"
repeat_region 4689. .4855
/rpt_family="B4"
repeat_region 5098. .5221
/rpt_family="L1"
repeat_region 5661. .5767
/rpt_family="Alu"

```

```

repeat_region      8126..8198
                    /rpt_family="ID"
repeat_region      8204..8266
                    /rpt_family="MIR"
repeat_region      8411..8661
                    /rpt_family="B4"
repeat_region      8662..8856
                    /rpt_family="B2"
                    complement(8778..8950)
trna                /product="tRNA-Ser"
                    /note="Likely pseudogene (MM Sc=38.89 / Sec struct
                    Sc=-12.60)
repeat_region      9012..9089
                    /rpt_family="ID"
repeat_region      9267..9417
                    /rpt_family="Alu"
repeat_region      10811..10955
                    /rpt_family="Alu"
repeat_region      11966..12120
                    /rpt_family="L1"
repeat_region      13709..14011
                    /rpt_family="L2"
repeat_region      14041..14240
                    /rpt_family="B2"
repeat_region      14351..14407
                    /rpt_family="Alu"
repeat_region      14444..14593
                    /rpt_family="B4"
repeat_region      14621..14752
                    /rpt_family="B4"
repeat_region      14693..14814
                    /rpt_family="Alu"
repeat_region      15214..15344
                    /rpt_family="MER1_type"
repeat_region      15710..16062
                    /rpt_family="L1"
repeat_region      16158..16300
                    /rpt_family="Alu"
repeat_region      16387..16791
                    /rpt_family="L1"
repeat_region      16912..17046
                    /rpt_family="L1"
repeat_region      17125..17225
                    /rpt_family="L1"
repeat_region      21101..21329
                    /rpt_family="L2"
repeat_region      21745..21839
                    /rpt_family="B4"
repeat_region      22054..22270
                    /rpt_family="L2"
repeat_region      22424..22598
                    /rpt_family="B2"
repeat_region      22746..23120
                    /rpt_family="L1"
repeat_region      23511..23650
                    /rpt_family="B2"
repeat_region      24060..24261
                    /rpt_family="MALR"
repeat_region      24265..24399
                    /rpt_family="MALR"
repeat_region      24400..24669
                    /rpt_family="B4"
repeat_region      24670..24828
                    /rpt_family="MALR"
repeat_region      24906..25850
                    /rpt_family="MALR"
repeat_region      26212..26318
                    /rpt_family="B2"
repeat_region      28076..28169
                    /rpt_family="MALR"
unsure             28600..28626
                    /note="Sequence derived from one plasmid subclone."
repeat_region      30605..30660

```

```

repeat_region      /rpt_family="ID"
                    30903..31077
repeat_region      /rpt_family="L1"
                    31339..31565
repeat_region      /rpt_family="B4"
                    31585..31631
repeat_region      /rpt_family="ERV1"
                    31709..31791
repeat_region      /rpt_family="B4"
                    32334..32483
repeat_region      /rpt_family="Alu"
                    32610..32756
repeat_region      /rpt_family="B4"
                    33052..33360
repeat_region      /rpt_family="L1"
                    33997..34123
repeat_region      /rpt_family="L1"
                    35091..35172
repeat_region      /rpt_family="L1"
                    35352..35526
repeat_region      /rpt_family="B2"
                    35583..35696

```

## Alignment Scores:

```

Pred. No.:      292      Length:      188476
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      6.34%      Indels: 0
DB:              10      Gaps: 0

```

US-09-690-825-34 (1-142) x AC122238 (1-188476)

QY 82 SerGlyCysAlaPheLeuSerValIys 90

DB 37686 TCTGGGTGCGCCTCTCTCAGTAAAG 37712

## RESULT 73

AL607066

LOCUS

DEFINITION Mouse DNA sequence from clone RP23-406N5 on chromosome 4, complete sequence.

ACCESSION AL607066

VERSION AL607066.23 GI:22213654

KEYWORDS HTG.

SOURCE Mus musculus

ORGANISM Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 194837)

REFERENCE

AUTHORS Lloyd,D.

TITLE Direct Submission

JOURNAL Submitted (09-AUG-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

On Aug 13, 2002 this sequence version replaced gi:21953272.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate accession numbers given in the feature table with their source databases: Emi, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-406N5 is from the RPCR-23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.choiri.org/bacpac/home.htm>

FEATURES  
source  
1. 194837  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="4"  
/clone="RP23-406N5"  
/clone\_1lb="RPCR-23"

## ORIGIN

## Alignment Scores:

Score:	Length:	Matches:	Mismatches:	Indels:	Gaps:
302	194837	9	0	0	0
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	6.34%				

US-09-690-825-34 (1-142) x AL607066 (1-194837)

Qy 133 ArgAlaIleGluInLeuAlaAlaMet 141

Db 131454 AGAGCAATTGACGACATGACGACATG 11480

RESULT 74  
AC109161/c

LOCUS  
DEFINITION  
MUS musculus clone RP24-112H1, WORKING DRAFT SEQUENCE, 20 ordered pieces.

AC109161  
AC109161.4 GI:11442509  
HTG; HTGS\_PHRASE2; HTGS\_DRAFT; HTGS\_FULFILLTOP.  
MUS musculus (house mouse)

## ORGANISM

MUS musculus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 199590)  
Britten, B., Nussbaum, C. and Lander, E.  
MUS musculus, clone RP24-112H1  
Unpublished

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCES  
AUTHORS

2 (bases 1 to 199590)  
Britten, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, W., Collins, S., Collamore, A., Cook, A., Cooke, P., Dekrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Gande, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneses, L., Minova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Ribbeck, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strause, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Trivis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,

TITLE  
JOURNAL  
RESEARCH  
320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 199590)  
Britten, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., Dekrellano, K., Collamore, A., Cook, A., Dooley, K., Dorris, L., Erickson, J., Fato, S., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneses, L., Minova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, V., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataratnam, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, D., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT  
Submitted (06-JUN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 6, 2003 this sequence version replaced GI:20303722.  
All repeats were identified using RepeatMasker:  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIRB

Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

Project Information  
Center project name: L18917  
Center clone name: 112\_H-1

Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 194743 bases at least Q40  
Consensus quality: 196079 bases at least Q20  
Consensus quality: 197024 bases at least Q20  
Insert size: 183000; agarose-fp  
Insert size: 197690; sum-of-contigs  
Quality coverage: 12.2 in Q20 bases; agarose-fp  
Quality coverage: 11.3 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 20 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.  
1  
18112: contig of 18112 bp in length  
18113: gap of 100 bp  
18114: contig of 411 bp in length  
18115: gap of 100 bp  
18116: contig of 100 bp  
18117: gap of 100 bp  
18118: contig of 766 bp in length  
18119: gap of 100 bp  
18120: contig of 582 bp in length  
18121: gap of 100 bp  
18122: contig of 706 bp in length  
18123: gap of 100 bp  
18124: contig of 685 bp in length  
18125: gap of 100 bp  
18126: contig of 100 bp

```

* 21863 22648: contig of 786 bp in length
* 22649 22748: gap of 100 bp
* 22749 23615: contig of 867 bp in length
* 23616 23715: gap of 100 bp
* 23716 24867: contig of 1152 bp in length
* 24868 24967: gap of 100 bp
* 24968 26657: contig of 1690 bp in length
* 26658 26757: gap of 100 bp
* 26758 28082: contig of 1325 bp in length
* 28083 28182: gap of 100 bp
* 28183 29563: contig of 1381 bp in length
* 29564 31551: contig of 1888 bp in length
* 31552 31651: gap of 100 bp
* 31652 34091: contig of 2440 bp in length
* 34092 34191: gap of 100 bp
* 34192 39340: contig of 5149 bp in length
* 39341 39440: gap of 100 bp
* 39441 44766: contig of 5326 bp in length
* 44767 44866: gap of 100 bp
* 44867 71940: contig of 27074 bp in length
* 71941 72040: gap of 100 bp
* 72041 98326: contig of 26286 bp in length
* 98327 98426: gap of 100 bp
* 98427 169712: contig of 71286 bp in length
* 169713 169812: gap of 100 bp
* 169813 199590: contig of 23778 bp in length.

FEATURES
    source
        1..199590
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="RP24-112H1"
            /clone_lib="RPCT-24 Male Mouse BAC"
            1..18112
                /note="assembly_fragment"
                clone_end:SP6
                vector_side:left"
            18213..18623
                /note="assembly_fragment"
            18724..19489
                /note="assembly_fragment"
            19590..20171
                /note="assembly_fragment"
            20272..20977
                /note="assembly_fragment"
            21078..21762
                /note="assembly_fragment"
            21863..22648
                /note="assembly_fragment"
            22749..23615
                /note="assembly_fragment"
            23716..24867
                /note="assembly_fragment"
            24968..26657
                /note="assembly_fragment"
            26758..28082
                /note="assembly_fragment"
            28183..29563
                /note="assembly_fragment"
            29664..31551
                /note="assembly_fragment"
            31652..34091
                /note="assembly_fragment"
            34192..39340
                /note="assembly_fragment"
            39441..44766
                /note="assembly_fragment"
            44867..71940
                /note="assembly_fragment"
            72041..98326
                /note="assembly_fragment"
            98427..169712
                /note="assembly_fragment"

```

```

misc_feature /note="assembly_fragment"
169813..199590
/note="assembly_fragment"
clone_end:T7
vector_side:right"

ORIGIN
Alignment Scores:
Pred. No.: 309 Length: 199590
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.34% Indels: 0
DB: 2 Gaps: 0

US-09-690-825-34 (1-142) x AC109161 (1-199590)
Qy 118 AsnAsnVslvslvsglupheglucgu 126
Db 47550 AACATPAAAGAAAGATTGAAGAG 47524

RESULT 75
AL805939 202366 bp DNA linear HTG 06-JUL-2002
LOCUS Mus musculus chromosome 4 clone RP23-327K19, 8 unordered pieces.
DEFINITION AL805939
ACCESSION AL805939
VERSION AL805939.8 GI:21711913
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 202366)

REFERENCE
AUTHORS Plumb,B.
TITLES Direct Submission
JOURNAL Submitted (04-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Jul 8, 2002 this sequence version replaced gi:21684715.

COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: BM327K19
----- Summary Statistics
Assembly program: XGAP4, version 4.5
Chemistry: Dye-terminator, 100% of reads
Consensus quality: 200675 bases at least Q40
Consensus quality: 201042 bases at least Q30
Consensus quality: 201317 bases at least Q20
Insert size: 201666; sum-of-contigs
Insert size: 202485; 5.9% error; agarose-fp
Quality coverage: 7.92x in Q20 bases; sum-of-contigs Quality
coverage: 8.00x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 31397: contig of 31397 bp in length
* 31398 31497: gap of 100 bp
* 31498 37803: contig of 6306 bp in length
* 37804 37903: gap of 100 bp
* 37904 98452: contig of 57549 bp in length
* 98453 98552: gap of 100 bp
* 98553 100802: contig of 5250 bp in length
* 100803 100902: gap of 100 bp

```

```

* 100903 114281: contig of 13379 bp in length
* 114382 114381: gap of 100 bp
* 114382 183917: contig of 69536 bp in length
* 183918 184017: gap of 100 bp
* 184018 188034: contig of 4017 bp in length
* 188035 188134: gap of 100 bp
* 188135 202366: contig of 14232 bp in length.

```

## FEATURES

```

1..202366
location/Qualifiers

```

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"

```

```

/clone="RP23-327K19"
/clone_1b="RPCI-23"

```

```

1..33397

```

```

/note="assembly_fragment:01971
clone_end:T7
vector_side:left"

```

```

31498..37803

```

```

/note="assembly_fragment:00959"
37904..95452

```

```

/note="assembly_fragment:01478
fragment_chain:1"

```

```

95553..100802

```

```

/note="assembly_fragment:00423
fragment_chain:1"

```

```

100903..114281

```

```

/note="assembly_fragment:01472
fragment_chain:1"

```

```

114382..183917

```

```

/note="assembly_fragment:01671
fragment_chain:1"

```

```

184018..188034

```

```

/note="assembly_fragment:02679
fragment_chain:1"

```

```

188135..202366
/note="assembly_fragment:00191
fragment_chain:1
clone_end:SP6
vector_side:right"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.:

```

```

Score: 313

```

```

Percent Similarity: 9.00

```

```

Best Local Similarity: 100.00%

```

```

Query Match: 6.34%

```

```

DB: 2

```

```

US-09-690-825-34 (1-142) x AL805939 (1-202366)

```

```

QY 133 ArgAlaIleGInLeuAlaIaMet 141

```

```

Db 158009 AGAGCAATTGACAACTAGCAGCATG 158035

```

```

Search completed: August 16, 2004, 02:46:59

```

```

Job time : 4302 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 16, 2004, 01:12:49 / Search time 2761 Seconds  
(without alignments)  
1535.831 Million cell updates/sec

Title: US-09-690-825-34  
Perfect score: 1  
Sequence: 1 MGAPPLPAPWOPFLKDHRS.....EFETAKKVRRALEQLAAMD 142

Scoring table:  
OLIGO  
Xgapop 60.0 / Xgapext 60.0  
Ygapop 60.0 / Ygapext 60.0  
Fgapop 6.0 / Fgapext 7.0  
Delop 6.0 / Delext 7.0

Searched: 27513289 segs, 14931090276 residues

Word size: 1

Total number of hits satisfying chosen parameters: 55025471

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgm2\_1/USPRO.spool/US09690825/runat\_11082004\_141015\_13804/app\_query.fasta\_1.327  
-DB=EST -OPMT=fastap -SUPPLX=oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100  
-DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEARST=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09690825.QCGN\_1\_1\_3437@runat\_11082004\_141015\_13804 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :  
EST :  
1: em\_estba :  
2: em\_esthum :  
3: em\_estlin :  
4: em\_estmu :  
5: em\_estov :  
6: em\_estpl :  
7: em\_estro :  
8: em\_hnc :  
9: gb\_hnc :  
10: gb\_est1 :  
11: gb\_hnc :  
12: gb\_est3 :  
13: gb\_est4 :  
14: gb\_est5 :  
15: em\_esthum :  
16: em\_estom :  
17: em\_gss\_hum :  
18: em\_gss\_inv :  
19: em\_gss\_pln :  
20: em\_gss\_vrt :  
21: em\_gss\_fun :  
22: em\_gss\_mam :  
23: em\_gss\_mus :  
24: em\_gss\_pro :  
25: em\_gss\_rtd :  
26: em\_gss\_phg :  
27: em\_gss\_vrl :  
28: gb\_gss1 :

29: gb\_gss2 :  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query Match	length	DB ID	Description
1	142	100.0	472	12	BM754864
2	142	100.0	478	9	BM710957
3	142	100.0	517	10	AM411195
4	142	100.0	552	12	BG258128
5	142	100.0	552	12	BM839824
6	142	100.0	553	14	CB138359
7	142	100.0	563	12	BM835829
8	142	100.0	574	14	CB111125
9	142	100.0	580	10	AM409701
10	142	100.0	584	14	CB128010
11	142	100.0	585	12	BG776624
12	142	100.0	602	9	AL603004
13	142	100.0	605	12	BM841904
14	142	100.0	613	10	AM409816
15	142	100.0	624	14	CB135918
16	142	100.0	629	10	BE292733
17	142	100.0	644	12	BM788077
18	142	100.0	646	12	B1252508
19	142	100.0	682	10	BE796084
20	142	100.0	698	10	BE796061
21	142	100.0	703	10	BE719993
22	142	100.0	707	12	BE115707
23	142	100.0	718	12	BM827836
24	142	100.0	726	13	BM421971
25	142	100.0	742	10	BE907403
26	142	100.0	760	12	BM020628
27	142	100.0	766	12	BM682887
28	142	100.0	768	12	B1771953
29	142	100.0	772	12	BE750227
30	142	100.0	785	13	BM600185
31	142	100.0	815	12	BM335442
32	142	100.0	815	12	BM744230
33	142	100.0	821	12	BM542207
34	142	100.0	822	9	AU130493
35	142	100.0	826	13	BUT171608
36	142	100.0	828	13	BUT171608
37	142	100.0	836	9	AU126048
38	142	100.0	861	13	BQ438198
39	142	100.0	868	13	BQ882604
40	142	100.0	874	10	BE883968
41	142	100.0	875	10	BE898710
42	142	100.0	875	12	BG528273
43	142	100.0	880	12	BM450871
44	142	100.0	883	13	BQ650551
45	142	100.0	888	13	BQ227378
46	142	100.0	889	12	BG250436
47	142	100.0	896	13	BK425472
48	142	100.0	908	10	BF972840
49	142	100.0	915	12	BG110508
50	142	100.0	924	12	B1760504
51	142	100.0	933	13	BQ955492
52	142	100.0	934	10	BF982118
53	142	100.0	935	10	BE883927
54	142	100.0	943	10	BF981768
55	142	100.0	943	13	BQ929774
56	142	100.0	962	13	BM649501
57	142	100.0	963	12	B1114304
58	142	100.0	965	13	BUT501647
59	142	100.0	974	12	BM463502
60	142	100.0	976	13	BQ064557
61	142	100.0	985	12	BM802009
62	142	100.0	990	13	BQ958253

```
63 142 100.0 993 12 BM564739
64 142 100.0 995 12 BM471181
65 142 100.0 1009 9 AL545180
66 142 100.0 1031 9 BF033760
67 142 100.0 1035 12 BM552163
68 142 100.0 1047 12 BQ048975
69 142 100.0 1051 12 BG035337
70 142 100.0 1052 12 BQ035337
71 142 100.0 1071 13 BQ073169
72 142 100.0 1073 13 BQ394560
73 142 100.0 1114 12 BM551159
74 142 100.0 1115 9 BQ053391
75 142 100.0 1115 9 AL549362
76 142 100.0 1201 9 AL515837
77 142 100.0 1201 9 AL553525
78 142 100.0 1201 13 BX416171
79 142 100.0 1201 13 BX422525
80 141 99.3 459 12 BG472824
81 141 99.3 1037 12 BM452294
82 141 99.3 1124 10 BE794850
83 140 98.6 712 12 BG830958
84 140 98.6 842 14 BE867588
85 139 97.9 462 14 CB136873
86 139 97.9 1055 12 BQ890576
87 139 97.9 1139 12 BM549522
88 138 97.2 532 10 BE892608
89 138 97.2 535 13 BU659089
90 138 97.2 575 10 BE786028
91 138 97.2 584 10 BE269775
92 133 93.7 757 12 BT334392
93 131 92.3 1156 12 BM465375
94 130 91.5 599 10 BE615441
95 129 90.8 718 10 BF219949
96 129 90.8 882 12 BT114918
97 129 90.8 923 12 BG748549
98 129 90.8 1056 10 BE686145
99 128 90.1 557 12 BM838065
100 128 90.1 788 12 BG254935
```

## ALIGNMENTS

```
RESULT 1
LOCUS BM754864
DEFINITION K-EST0032467 S11SNUI Homo sapiens cDNA clone S11SNUI-12-C02 5',
mRNA sequence.
```

```
ACCESSION BM754864
VERSION BM754864.1
KEYWORDS GI:19084482
EST.
```

```
SOURCE Homo sapiens (human)
ORGANISM
```

```
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 472)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
```

```
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 12 row: C column: 02
High quality sequence stop: 472.
```

```
FEATURES
source
```

```
1. 472
/organism="Homo sapiens"
/mol_type="mRNA"
```

```
/db xref="taxon:9606"
/clone="S11SNUI-12-C02"
/sex="M"
/tissue_type="Stomach"
/cell_line="SNU-1"
/clone_lib="S11SNUI"
/note="Organ: Stomach; Vector: pME18-PJ3; Site 1: XhoI;
Site 2: XhoI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deacapped
with labacco acid pyrophosphatase (TAP). The deacapped
intracl mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfiI and
cloned into DraIII- digested pME18-PJ3 vector. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
```

## ORIGIN

```
Alignment Scores:
Pred. No.: 7,37e-149 Length: 472
Score: 142.00 Consesrative: 142
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 12 Gaps: 0
```

```
US-09-690-825-34 (1-142) x BM754864 (1-472)
```

```
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuYAspHisArgIleSer 20
DB 45 ATGGGTCGCCCAACGTCGTCCTGCGAGGCTGCGCTGACCCCGAGGGAGTGGCCGAG 104
QY 21 ThrPheLeuAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 105 ACATTCAAGAACTGGCGCTTCTTGAGAGGCTGCGCTGACCCCGAGGGAGTGGCCGAG 164
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGluCysPhePheCys 60
DB 165 GCTGGCTTATCATCAGTCCGCCCTGAGACAGCCAGCACTGGCCAGCTTCTTCTGCG 224
QY 61 PheLeuGluLeuGluGlyTrpGluProAspAspProIleGluGluHisIleCysIle 80
DB 225 TTCAAGAGAGCTGGAAGGCTGGAGCCAGTACGACCCCATAGAGGAAACATATAAGCAT 284
QY 81 SerSerGlyCysAlaPheLeuSerValIleGluGluGluGluGluGluGluGluGlu 100
DB 285 TCGTCGGTGTGGCTTCTTCTTCTGCAAGAGAGGAGTGAAGAAATTAACCTTGCTGAA 344
QY 101 PheLeuLeuLeuAspArgGluArgAlaIleAsnIleAlaIleAlaIleAlaIleAla 120
DB 345 TTTTGAAGACTGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 404
QY 121 LysGlyGluPheGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 140
DB 405 AAGAAAGAAATTGAGAGAACTGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 464
QY 141 MetAsp 142
DB 465 ATTCGAT 470
```

```
RESULT 2
LOCUS AL710957
DEFINITION DKF20686A0177.r1.686 (synonym: h1cc3) Homo sapiens cDNA clone
ACCESSION AL710957
```



```

VERSION      AT1710957.1  GI:19694312
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE       1 (bases 1 to 478)
JOURNAL      Wamburt,R., Heubner,D., Wewes,W., Weill,B. and Wiemann,S.
COMMENT      EST (Wamburt,R., Heubner,D., Wewes,H.W., Weill,B. and Wiemann,S.)
              Unpublished (1999)
              Contact: MIPS
              MIPS
              Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
              This is the 5' sequence of the clone insert
              Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
              Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
              sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
              consortium of the German Genome Project.
              No s1 sequence available.
              This clone (DKFZ686A0177) is available at the RZPD in Berlin.
              Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
              Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
  source
    1..478
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="DKFZ686A0177"
      /dev_stage="adult"
      /lab_host="DH10B"
      /clone_lib="686 (synonym: hicc3)"
      /note="Vector: pTR10Ex2; Site_1: sf11A; site_2: sf11B;
              cDNA-collection"
ORIGIN
Alignment Scores:
  Pred. No.:      7,45e-149      Length:      478
  Score:          142.00         Matches:      142
  Percent Similarity: 100.00%    Conservative: 0
  Best Local Similarity: 100.00% Mismatches:    0
  Query Match:    100.00%       Indels:        0
  DB:              9            Gaps:          0
US-09-690-825-34 (1-142) x AT1710957 (1-478)
QY      1 MetG1yA1aPrOthLeuProFroAlaTrpGlnProPhelLeuYsAspH1sarG1leSer 20
Db      22 ATGGGTGCCCCGAGCTTGGCCCTGCGAGCCCTTCTCAAGACCAACCGATCTCT 81
QY      21 ThirPhelYsAsnTrpProPhelLeuGluGlyCysAlaCysThrProGluUArgMetAlaGlu 40
Db      82 ACATTCAGAACTGCGCCCTTCTGAGGGCTGCGCCCTGCACCCCGAGCGGATGGCCGAG 141
QY      41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db      142 GCTGGCTTCACTCCACCTGCCCCACCTGAGACGACCACTTGGCCCGATGTTCTTCTGCG 201
QY      61 PhelYsGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisIleYsHis 80
Db      202 TTCAGAGAGCTGGAAGCTGGAGCCAGATGACCAACCCCATAGAGGAACATAAAAGCAT 261
QY      81 SerSerGlyCysAlaPheLeuSerValYsYsGlnPheGluGluLeuThrLeuGlu 100
Db      262 TCGTCCGGTTCCTTCTTCTCTGTCAGAGAGAGATTGAAGATTAAACCTTGGTGA 321
QY      101 PheLeuYsLeuAspArgGluArgAlaYsAsnYsIleAlaYsGluThrAsnAsnYs 120
Db      322 TTTTGAATCTGACAGAGAAAGAGCAACAATTTGCAAGAGAAACCAACATTAAG 381
QY      121 YsYsGluPheGluGluGluThrAlaYsYsValArgArgAlaIleGluGluAla 140
Db      382 AAGAAAGATTGAGAAATGCGAAGAAAGTGGCCGTCATCGAGCAGCTGGCTCC 441

```

```

QY      141 MetAsp 142
Db      442 ATGGAT 447
RESULT 3
AM411195
LOCUS     517 bp      mRNA      linear      EST 29-JUN-2000
DEFINITION fthcd09.x1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2964713 5',
            mRNA sequence.
ACCESSION AM411195
VERSION   AM411195
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE     1 (bases 1 to 517)
JOURNAL   NIH-MGC http://mgs.nci.nih.gov/.
COMMENT   Unpublished (1999)
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Contact: Robert Strausberg, Ph.D.
           Email: cgapsb-remail.nih.gov
           Tissue Procurement: ATCC
           cDNA Library Preparation: Ling Hong/Rubin Laboratory
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
           DNA Sequencing by: National Institutes of Health Intramural
           Sequencing Center (NISC)
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LNLN at:
           www-bio.lnl.gov/btrp/image/image.html
           Plate: LHC64 row: 0 column: 18
           Seg primer: -21M13 forward primer (ABI).
FEATURES
  source
    1..517
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:2964713"
      /tissue_type="Thadomyosarcoma"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH MGC 17"
      /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
              Site_2: XhoI; cDNA made by c1igo-dt priming.
              Directionally cloned into EcoRI/XhoI sites using the
              following 5' adaptor: GGACACAG(G). Size-selected >500bp
              for average insert size 1.8kb. Library constructed by
              Ling Hong in the laboratory of Gerald M. Rubin (University
              of California, Berkeley) using ZAP-cDNA synthesis kit
              (Stratagene) and Superscript II RT (Life Technologies). "
ORIGIN
Alignment Scores:
  Pred. No.:      7,96e-149      Length:      517
  Score:          142.00         Matches:      142
  Percent Similarity: 100.00%    Conservative: 0
  Best Local Similarity: 100.00% Mismatches:    0
  Query Match:    100.00%       Indels:        0
  DB:              10            Gaps:          0
US-09-690-825-34 (1-142) x AM411195 (1-517)
QY      1 MetG1yA1aPrOthLeuProFroAlaTrpGlnProPhelLeuYsAspH1sarG1leSer 20
Db      44 ATGGGTGCCCCGAGCTTGGCCCTGCGAGCCCTTCTCAAGACCAACCGATCTCT 103
QY      21 ThirPhelYsAsnTrpProPhelLeuGluGlyCysAlaCysThrProGluUArgMetAlaGlu 40
Db      104 ACATTCAGAACTGCGCCCTTCTGAGGGCTGCGCCCTGCACCCCGAGCGGATGGCCGAG 163
QY      41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db      164 GCTGGCTTCACTCCACCTGCCCCACCTGAGACGACCACTTGGCCCATGTTCTTCTGCG 223

```

```

QY      61  PheLysGluLeuGluGlyTrrpGluProAspAspProIleGluGluHis 80
DB      224 TTCAAGAGAGCTGAGAGCGTGGAGCCATGACGACCCCATAGAGAACTAAACAT 283
QY      81  SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB      284 TCGTCGGGTTGGCTTCTTCTTCTGCAAGAGCGATTGAAGATTAACCTTGGGAA 343
QY      101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
DB      344 TTTTGAAGACTGAGAGAGAAAGCCAGAACTTGCAGAGAGAAACCAACATAG 403
QY      121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaAla 140
DB      404 AAGAAGAAATTGAGAGAACTCGAAGAAAGTCCCTGCATCGAGCACTGCTGCC 463
QY      141 MetAsp 142
DB      464 ATGGAT 469

RESULT 4
BG258128      552 bp      mRNA      linear      EST 13-FEB-2001
LOCUS      602379226f1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4510014 5'
DEFINITION      mRNA sequence.
ACCESSION      BG258128
VERSION      BG258128.1 GI:1267944
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 552)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: LHAM10391 row: C column: 07
              High quality sequence stop: 552.
FEATURES
    source
        1..552
            Location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4510014"
                /tissue_type="embryonal carcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH MGC 92"
                /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NCI;
                Site 2: Sall; Cloned unidirectionally; oligo-dT primed.
                Average insert size 2.5 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.:      8-42e-149      Length:      552
Score:          142.00      Matches:      142
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:            12      Gaps:      0
US-09-690-825-34 (1-142) x BG258128 (1-552)

```

```

QY      1  MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuAspHisArgIleSer 20
DB      32  ATGGAGGCCCGACGTTGCCCCCTGCGACGCCCTTCTCAAGAGACACCGCATCTCT 91
QY      21  ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB      92  ACATTCAAACTGCGCTTCTTCTTGAAGGCTGCGCTGACCCCGAGCGAGTGGCCAG 151
QY      41  AlaGlyPheIleHisCysProThrGluGlnGluProAspLeuAlaGlnCysPhePheCys 60
DB      152 GCTGGCTTATCATCACTGCCCTGAGAGAGAGCCAGACTTGGCCAGATTCTTCTGCG 211
QY      61  PheLysGluLeuGluGlyTrrpGluProAspAspProIleGluGluHisLysLysHis 80
DB      212 TTCAAGAGACTGAGAGCGTGGAGGCAAGTGAAGACCCCATAGAGAGAACTAAAAAGAT 271
QY      81  SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB      272 TCGTCGGGTTGGCTTCTTCTTCTGCAAGAGCGATTGAAGATTAACCTTGGTAA 331
QY      101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
DB      332 TTTTGAAGACTGAGAGAGAAAGCCAGAACTTGCAGAGAGAAACCAACATAG 391
QY      121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaAla 140
DB      392 AAGAAGAAATTGAGAGAACTCGAAGAAAGTCCCTGCATCGAGCACTGCTGCC 451
QY      141 MetAsp 142
DB      452 ATGGAT 457

RESULT 5
BM839824      552 bp      mRNA      linear      EST 06-MAR-2002
LOCUS      K-EST0116794 S13KMS5 Homo sapiens cDNA clone S13KMS5-32-F06 5'
DEFINITION      mRNA sequence.
ACCESSION      BM839824
VERSION      BM839824.1 GI:19196233
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 552)
AUTHORS      Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
              Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
              Kim,Y.S.
TITLE      21C Frontier Korean EST Project 2001
JOURNAL      Unpublished (2002)
COMMENT      Contact: Kim YS
              Genome Research Center
              Korea Research Institute of Bioscience & Biotechnology
              52 Beon-dong Yuseong-gu, Daejeon 305-353, South Korea
              Tel: +82-42-860-4470
              Fax: +82-42-860-4409
              Email: yongsung@mail.kribb.re.kr
              Place: 32 row: F column: 06
              High quality sequence stop: 552.
FEATURES
    source
        1..552
            Location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="S13KMS5-32-F06"
                /tissue_type="myeloma"
                /cell_line="KMS-5"
                /lab_host="Hsp10P"
                /clone_lib="S13KMS5"
                /note="Vector: pCNS; Site 1: EcoRI; Site 2: NotI; The poly
                (A) + RNA was dephosphorylated with bacterial alkaline
                phosphatase (BAP) and then decapped with tobacco acid
                pyrophosphatase (TAP). The decapped intact mRNA was

```

ligated with DNA-RNA linker including Ecor I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10<sup>®</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

## ORIGIN

## Alignment Scores:

Pred. No.:	8,42e-149	Length:	552
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-09-690-825-34 (1-142) x BM839824 (1-552)

```
QY      1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
DB      78 ATGGGTGCCCCGAGCTGGCCCCCTGGCCAGCCCTTCTCAAGGACCCGCACTCTCT 137
QY      21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB     138 ACATTCAAGAACTGCGCCCTTCTTGAGGCGCTGCGCCCTGCACCCCGAGCGCGAG 137
QY      41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB     198 GCTGCGCTTCACTCCCTCCCACTGAGAACGACGACCACTTGCGCCAGTCTTCTCTGC 257
QY      61 PheLysGluPheGluGluGlyTrpGluProAspAspProIleGluGluHis 80
DB     258 TTCAAGAGAGCTGGAGGCTGGGAGCCAGATGACAGCCCTTAGAGAACTATAAAGCAT 317
QY      81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100
DB     318 TCGCCGAGTCCGCTTCTTCTTCTGTCAGAGACAGTTTGAAGATTAACCTTGCTGAA 377
QY     101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
DB     378 TTTTGAAGACTGACAGAGAAAGACCAAAATTGCAAGAGAAACCAACAACTAAG 437
QY     121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140
DB     438 AAGAAAGATTTTGAAGAACTGCAAGAAAGAGTGGCGGTCATCGAGCAGCTGGCTGCC 497
QY     141 MetAsp 142
DB     498 ATGGAT 503

RESULT 6
CB138359      553 bp      mRNA      linear      EST 29-JAN-2003
LOCUS      K-BST0191218 LSHUK1 Homo sapiens cDNA clone LSHUK1-23-G07 5', mRNA
DEFINITION      Sequence.
ACCESSION      CB138359
VERSION      CB138359.1 GI:28109974
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 553)
AUTHORS      Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE      21c Frontier Korean EST Project 2001
```

JOURNAL  
COMMENT

Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 23 row: G column: 07  
High quality sequence stop: 553.  
Location/Qualifiers

## FEATURES

## source

1..553  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="LSHUK1-23-G07"  
/sex="M"  
/cell\_line="HLK-1"  
/lab\_host="Top10<sup>®</sup>"  
/clone\_lib="LSHUK1"  
/note="Organ: Liver; Vector: pcNS-D2; Site 1: EcoRI;  
Site 2: NotI; The poly (A) + RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then deacapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including  
EcoRI site by treatment of T4 RNA ligase and the first  
strand cDNA was synthesized from oligo dT-selected mRNA by  
priming with dT-tailed vector. The dT-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10<sup>®</sup> by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."

## ORIGIN

## Alignment Scores:

Pred. No.:	8,43e-149	Length:	553
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-09-690-825-34 (1-142) x CB138359 (1-553)

```
QY      1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
DB     40 ATGGGTGCCCCGAGCTGGCCCCCTGGCCAGCCCTTCTCAAGAGACCAAGCATCTCT 99
QY     21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB    100 ACATTCAAGAACTGCGCCCTTCTTGAGGCGCTGCGCCCTGCACCCCGAGCGAGTGCAG 159
QY     41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB    160 GCTGCGCTTCACTCCCTCCCACTGAGAACGACGACCTTGCCCAAGTCTTCTTGC 219
QY     61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80
DB    220 TTCAAGAGAGCTGGAAGGCTGGAGCGGATGACGACCCCATAGAGAACTATAAAGCAT 279
QY     81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluThrLeuGlyGlu 100
DB    280 TCGTCCGAGTTCGCTTCTTCTTCTGTCAGAGACGATTTGAAGATTAACCTTGCTGAA 339
QY     101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
DB    340 TTTTGAAGACTGACAGAGAAAGACCAAAATTGCAAGAGAAACCAACAACTAAG 399
QY     121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140
```

```

Db      400 AAGAAAGATTGAGGAACCTGCGAAGAACTGGCGCTGCCTCGAGACGCTGCTGCC 459
QY      141 Metasp 142
Db      460 ATGGAT 465

RESULT 7
LOCUS   BM835829
DEFINITION K-EST0111069 S11SN01 Homo sapiens cDNA clone S11SN01-73-H09 5',
            mRNA sequence.
ACCESSION BM835829
VERSION   BM835829.1 GI:19192238
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 563)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
COMMENT  Contact: Kim YS
        Genome Research Center
        Korea Research Institute of Bioscience & Biotechnology
        52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
        Tel: +82-42-860-4470
        Fax: +82-42-860-4409
        Email: yongsung@mail.kribb.re.kr
        Plate: 73 row: H column: 09
        High quality sequence stop: 563.
        Location/Qualifiers
            1..563
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="S11SN01-73-H09"
                /sex="M"
                /tissue_type="Stomach"
                /cell_type="Lymphoblast-like"
                /lab_host="Top10F"
                /clone_lib="S11SN01"
                /note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
                Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
                bacterial alkaline phosphatase (BAP) and then decapped
                with tobacco acid pyrophosphatase (TAP). The decapped
                intact mRNA was ligated with DNA-RNA linker including SfiI
                site by treatment of T4 RNA ligase and the first strand
                cDNA was synthesized with Superscript II using SfiI
                oligo-dT primer. After first strand synthesis, RNA was
                degraded by NaOH treatment and cDNA was amplified by PCR
                reaction. The PCR products were digested with SfiI and
                cloned into DraIII-digested pME18S-FL3 vector. The
                obtained cDNA vectors were used for transformation of
                competent cells E. coli Top10F by electroporation method.
                The cDNA libraries constructed by this method are
                full-length enriched cDNA library."

ORIGIN
Alignment Scores:
Pred. No.:      8,566-149      Length:      563
Score:          142.00         Matches:      142
Percent Similarity: 100.00%     Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match:    100.00%         Indels:      0
DB:             12             Gaps:        0

US-09-690-825-34 (1-142) x BM835829 (1-563)

```

```

QY      1 MetGlyAlaProThrLeuProProAlaTProInProPheLeuYasPHisArgIleSer 20
Db      45 ATGGGAGGCCCGACGCGTGGCCCCCTCGCTGGCAGCCCTTCTCAAGAGCACCGCATCTCT 104
QY      21 ThrPheLeuAsnTTPProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db      105 ACATTCGAAGACTGGCCCTCTTGGAGGAGCTGGCTGACCCCGAGCGGAGTGCCAG 164
QY      41 AlaGlyPheIleHisCysPProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
Db      165 GCTGGCTTATCATCATCGCCCGCCACTGAGAACGAGCCAGACTGGCCCACTGTTCTTCG 224
QY      61 PheLeuGluLeuGluGlyTTPGluProAspAspAspProIleGluGluHisLeuLysHis 80
Db      225 TTCAGAGAGACTGGAAGGCTGGAGGAGCAGATGACGCCCAATGAGAAACATAAAAAGCAT 284
QY      81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
Db      285 TGCTCGGCTGGCGCTTCTTCTCTGCAAGAGCAGTTCGAAGATTACCTCTTGCA 344
QY      101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
Db      345 TTTTGAAACTGACAGAGAAAGAGCCCAAGAACAAATTTGAAAGAGAAACCAACATTAAG 404
QY      121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaAla 140
Db      405 AAGAAAGATTGAGGAACCTGCGAAGAACTGGCGCTGCCTCGAGACGCTGCTGCC 464
QY      141 Metasp 142
Db      465 ATGGAT 470

RESULT 8
LOCUS   CB111125
DEFINITION K-EST0152661 LSHK1 Homo sapiens cDNA clone LSHK1-8-G05 5', mRNA
            sequence.
ACCESSION CB111125
VERSION   CB111125
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 574)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
COMMENT  Contact: Kim YS
        Genome Research Center
        Korea Research Institute of Bioscience & Biotechnology
        52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
        Tel: +82-42-860-4470
        Fax: +82-42-860-4409
        Email: yongsung@mail.kribb.re.kr
        Plate: 8 row: G column: 05
        High quality sequence stop: 574.
        Location/Qualifiers
            1..574
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="LSHK1-8-G05"
                /sex="M"
                /cell_line="HLK-1"
                /lab_host="Top10F"
                /clone_lib="LSHK1"
                /note="Organ: Liver; Vector: pCNS-D2; Site_1: EcoRI;
                Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                bacterial alkaline phosphatase (BAP) and then decapped
                with tobacco acid pyrophosphatase (TAP). The decapped

```

intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10<sup>®</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

## ORIGIN

## Alignment Scores:

Pred. No.:	8,71e-149	Length:	574
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-09-690-825-34 (1-142) x CB111125 (1-574)

```
QY 1 MetGlyAlaProThrLeuProProAlaTrpGluProPheLeuLysAspHisArgIleSer 20
DB 55 ATGGGTGCCCCGAGCGTTGCCCCCTGCTGGCAGCCCTTCTCAAGACACCGCATCTCT 114
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 115 ACATTCAAGACTGCGCCCTTCTTGAGAGCGCTGGCCCTGGCAACCCCGAGCGGATGGCCGAG 174
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB 175 GCTGGCTTCATCCACTGCCCTCCACGAGACGAGCAACTGGCCCGCTTCTTCTTCTGC 234
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisIleSlySlyHis 80
DB 235 TTCAAGAGAGCTGAGAGCTGGAGCCAGATGACACCCCTTAGAGACATTAAAAAGCAT 254
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB 295 TCGCGCGGTGGCGCTTCTTCTGTCAGAGAGAGATTGAAGAAATTAAACCTTGTGAA 354
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
DB 355 TTTTGAACCTGACACAGAAAGAGCCAGAACAAATTGCAANGAGAAACCAACAAATAAG 414
QY 121 LysLysGluPheGluGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaAla 140
DB 415 AAGAAAGAAATTGAGAAAGCTGCGAAGAAAGTGGCCGCTCCATCGACGAGCTGGTGC 474
QY 141 MetAsp 142
DB 475 ATGGAT 480
```

RESULT 9  
AM409701 580 bp mRNA linear EST 29-JUN-2000  
LOCUS fl01e02.x1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:2961098 5'  
DEFINITION mRNA sequence.  
ACCESSION AM409701  
VERSION AM409701.1 GI:6935175  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 580)  
AUTHORS NIH-MGC <http://mgc.ncbi.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

## FEATURES

## source

Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
[www.bio.lnl.gov/bbrp/image/image.html](http://www.bio.lnl.gov/bbrp/image/image.html)  
Plate: LNCM55 row: I column: 3  
Seq primer: -21M3 forward primer (ABI).  
Location/Qualifiers  
1..580  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2961098"  
/tissue\_type="rhabdomyosarcoma"  
/lab\_host="MDH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_17"  
/note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;  
Site 2: XhoI; cDNA made by oligo-dt priming.  
directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

## Alignment Scores:

Pred. No.:	8,78e-149	Length:	580
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-690-825-34 (1-142) x AM409701 (1-580)

```
QY 1 MetGlyAlaProThrLeuProProAlaTrpGluProPheLeuLysAspHisArgIleSer 20
DB 53 ATGGGTGCCCCGAGCGTTGCCCCCTGCTGGCAGCCCTTCTCAAGACACCGCATCTCT 112
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 113 ACATTCAAGACTGCGCCCTTCTTGAGAGCGCTGGCCCTTCACCCCGAGCGGATGGCCGAG 172
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB 173 GCTGGCTTCATCCACTGCCCTCCACGAGACGAGCAACTGGCCCGCATGTTCTTCTGC 232
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisIleSlySlyHis 80
DB 233 TTCAAGAGAGCTGAGAGCTGGAGCCAGATGACGACCCCATAGAGAAACATPAAAAAGCAT 292
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB 293 TCGTCGGGTGGCGCTTCTTCTGTCAGAGAGCGTTGAAGATTAAACCTTGTGAA 352
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
DB 353 TTTTGAACCTGACACAGAAAGAGCCAGAACAAATTGCAANGAGAAACCAACAAATAAG 412
QY 121 LysLysGluPheGluGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaAla 140
DB 413 AAGAAAGAAATTGAGAAAGCTGCGAAGAAAGTGGCCGCTCCATCGACGAGCTGGTGC 472
QY 141 MetAsp 142
DB 473 ATGGAT 478
```

RESULT 10

CB128010  
 LOCUS CB128010 584 bp mRNA linear EST 29-JAN-2003  
 DEFINITION K-EST0177364 C15NU17 Homo sapiens cDNA clone C15NU17-12-G02 5',  
 mRNA sequence.  
 ACCESSION CB128010  
 VERSION CB128010.1 GI:28090807  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 584)  
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.krrib.re.kr  
 Plate: 12 row: G column: 02  
 High quality sequence stop: 584.  
 Location/Qualifiers  
 1..584  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="C15NU17-12-G02"  
 /sex="F"  
 /tissue\_type="uterine"  
 /cell\_line="Epithelial"  
 /cell\_type="SNU-17"  
 /lab\_host="TOP10F"  
 /clone\_id="C15NU17"  
 /note="Organ: Cervix; Vector: pCNS-D2; Site 1: EcoRI;  
 Site 2: NotI; The poly (A) + RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then deapped  
 with tobacco acid pyrophosphatase (TAP). The deapped  
 intact mRNA was ligated with DNA-RNA linker including  
 EcoRI site by treatment of T4 RNA ligase and the first  
 strand cDNA was synthesized from oligo dt-selected mRNA by  
 priming with dt-tailed vector. The dt-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was  
 circularized with E. coli DNA ligase after digestion of  
 EcoRI which site is also included in vector. An RNA strand  
 converted to a DNA strand by Okayama-Berg method. The  
 obtained cDNA vectors were used for transformation of  
 competent cells E. coli TOP10F by electroporation method.  
 The cDNA libraries constructed by this method are  
 full-length enriched cDNA library."

ORIGIN  
 Alignment Scores: 8.84e-149 Length: 584  
 Prid. No.: 142.00 Matches: 142  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 14 Gaps: 0  
 DB: 14  
 US-09-690-825-34 (1-142) x CB128010 (1-584)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 55 ATGGGAGCCCGACCGTGGCCCTCTGGCGAGCCCTTTCACAGACACCGCATCTCT 114  
 QY 21 ThrPheLysAsnTrpProPheLeuGlnGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 115 ACATTCAAGAACTGGCCCTCTTGTGAGGAGCGCTGGCGCTGCACCCGAGCGSAGTGGCGAG 174

QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 Db 175 GCTGGCTTCATCATCTGCGCCCACTGAGAACAGCCAGCACTGGCCCACTGTTCTTCGCG 234  
 QY 61 PheLysGluLeuGlnGlyTrpGluProAspAspAppProlleGluGlnHisLysHis 80  
 Db 235 TTCAGAGAGCTGAGAGCGCTGGAGCGACATGACGCCCATATGAGSAAACATMAAGCAT 294  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGlnGluLeuThrLeuGlyGlu 100  
 Db 295 TCGTCCGGTGGCTTCTCTCTCTCAAGAGCGATTGAGAAATTAACCTTGGTGA 354  
 QY 101 PheLeuLysLeuAspArgGluLysAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 Db 355 TTTTGAAGAACTGAGAGAGAAAGACCAAGAACAAAATTCGAAAGAAACCAACATATAG 414  
 QY 121 LysLysGluPheGlnGluThrAlaLysLysValaGAGAlaIleGluGlnLeuAla 140  
 Db 415 AAGAAAGAAATTTGAGAAACTGCGAAGAAAGTGGCGCTGATCGAGCACTGGCTGCC 474  
 QY 141 MetAsp 142  
 Db 475 ATCGAT 480

RESULT 11  
 BG776624 585 bp mRNA linear EST 15-MAY-2001  
 LOCUS BG776624  
 DEFINITION mRNA sequence.  
 ACCESSION BG776624  
 VERSION BG776624.1 GI:14046941  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 585)  
 NIH-MGC http://mhc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgaabs@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNLI at:  
 http://image.llnl.gov  
 Plate: LLCM1662 row: d column: 19  
 High quality sequence stop: 585.  
 Location/Qualifiers  
 1..585  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4808802"  
 /tissue\_type="mucoepidermoid carcinoma"  
 /lab\_host="DH10B (TI phage-resistant)"  
 /clone\_id="NIH\_MGC\_59"  
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:  
 SfiI (ggcgccctcgcc); Site 2: SfiI (ggcgcatatggcc);  
 Double-stranded cDNA was prepared from cell line RNA. 5'  
 and 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-ATTCTAGAGCCGAGCGCGCGGATG-dt(30)-BN-3'  
 (where B = A, C, or G and N = A, C, G, or T). Average  
 insert size 1.65 kb (range 0.9-4.0 kb). 15/15 clones  
 contained inserts by PCR. This library was enriched for  
 full-length clones and was constructed by Clontech  
 Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC  
 library."

FEATURES  
 source

## ORIGIN

## Alignment Scores:

Pred. No.:	8.85e-149	Length:	585
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-09-690-825-34 (1-142) x BG776624 (1-585)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 46 ATGGGTGCCCCGAGCGTTGCCCTGGCTGGCAGCCCTTTCTCAAGGACCAAGCATCTCT 105  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 106 ACATTCAAGAACTGCGCCCTTCTTGAGAGGCTGCGCCCTGACCCCGGAGCGGATGGCCGAG 165  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 Db 166 GCTGGCTTCACTCCCTCCCTCCACAGAGAGCAAGCATTTGGCCCACTGTTCTTCTGCTGC 225  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisIleLysLysHis 80  
 Db 226 TTCAAGAGAGCTGAGAGGCTGGAGCCAGATGACCAAGCCATAGAGGAACTTAAAGCAT 285  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 286 TCGTCCGCTTCCCTGCTTCTTCTGTCAGAGAGAGATTTGAAGATTTAACCTTGGGAA 345  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 Db 346 TTTTGAAGACTGACAGAGAAAGAGCCAGAACTTGCAGAAAGCAATTAAG 405  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaAla 140  
 Db 406 AAGAAAGAAATTGAGAAACTGCGAAGAAAGTGCGCCGTCATGAGCAAGCTGGCTGCC 465  
 QY 141 MetAsp 142  
 Db 466 ATGGAT 471

## RESULT 12

AL603004 602 bp mRNA linear EST 04-SEP-2003  
 LOCUS DKFZP686K1620.r1.686 (synonym: hicc3) Homo sapiens cDNA clone  
 DEFINITION DKFZP686K1620.5', mRNA sequence.

ACCESSION AL603004  
 VERSION AL603004.1 GI:15166510  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 602)  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 EST (Wambutt, R., Heubner, D., Mewes, W., Weil, B. and Wiemann, S.)  
 EST (Wambutt, R., Heubner, D., Mewes, H.W., Weil, B. and Wiemann, S.)  
 Unpublished (1999)  
 CONTACT: MIPS

## COMMENT

MIPS

INGOLTAEDTER Landstr. 1, D-85764 Neuberg, Germany  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;  
 sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing  
 consortium of the German Genome Project.  
 No sl sequence available.

This clone (DKFZP686K1620) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

Location/Qualifiers

## source

1..602  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DKFZP686K1620"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="686 (synonym: hicc3)"  
 /note="Vector: pTripleX2; Site\_1: SfiIA; Site\_2: SfiIB;  
 cDNA-collection"

## ORIGIN

## Alignment Scores:

Pred. No.:	9.07e-149	Length:	602
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-690-825-34 (1-142) x AL603004 (1-602)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 45 ATGGGTGCCCCGAGCGTTGCCCTGGCTGGCAGCCCTTTCTCAAGGACCAAGCATCTCT 104  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 105 ACATTCAAGAACTGCGCCCTTCTTGAGAGGCTGCGCCCTGACCCCGGAGCGGATGGCCGAG 164  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 Db 165 GCTGGCTTCACTCCCTCCCTCCACAGAGAGCAAGCATTTGGCCCACTGTTCTTCTGCTGC 224  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisIleLysLysHis 80  
 Db 225 TTCAAGAGAGCTGAGAGGCTGGAGCCAGATGACCAAGCCATAGAGGAACTTAAAGCAT 284  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 285 TCGTCCGCTTCCCTGCTTCTTCTGTCAGAGAGAGATTTGAAGATTTAACCTTGGGAA 344  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 Db 345 TTTTGAAGACTGACAGAGAAAGAGCCAGAACTTGCAGAAAGCAATTAAG 404  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaAla 140  
 Db 405 AAGAAAGAAATTGAGAAACTGCGAAGAAAGTGCGCCGTCATGAGCAAGCTGGCTGCC 464  
 QY 141 MetAsp 142  
 Db 465 ATGGAT 470

## RESULT 13

BM841904 605 bp mRNA linear EST 06-MAR-2002  
 LOCUS K-EST0119223.S13KMS5 Homo sapiens cDNA clone S13KMS5-34-A12.5',  
 DEFINITION mRNA sequence.

ACCESSION BM841904  
 VERSION BM841904.1 GI:19198313  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 605)  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 EST (Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
 Kim, Y.S., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
 Kim, Y.S.)  
 21C Frontier Korean EST Project 2001  
 Unpublished (2002)  
 CONTACT: Kim YS

## COMMENT

Location/Qualifiers

Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong, Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 34 row: A column: 12  
High quality sequence stop: 605.  
Location/Qualifiers

1. 605  
/organism="Homo sapiens"  
/mol\_type="cDNA"  
/db\_xref="taxon:9606"  
/clone="S13KMS5-34-A12"  
/tissue\_type="myeloma"  
/cell\_line="KMS-5"  
/lab\_host="Top10F"  
/clone\_lib="S13KMS5"  
/note="Vector: pcms; Site 1: EcoRI; Site 2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then dephosphorylated with tobacco acid pyrophosphatase (TAP). The dephosphorylated intact RNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

## ORIGIN

Alignment Scores:  
Pred. No.: 9, 116-149 Length: 605  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-690-825-34 (1-142) x BK841904 (1-605)

QY 1 MetGlyAlaProThrLeuProProAlaTirpGlnProPheLeuYAspHisArgIleSer 20  
DB 78 ATGGGTGCCCCGAGCTTCCCTGCGAGCCCTTTCAGAGCAACCGCATCTCT 137  
QY 21 ThrPheYAsnTirpProPheLeuGluGlyCysAlaCysTirpProGluArgMetAlaGlu 40  
DB 138 ACATTCAAGAACTGGCCCTCTTGTGAGGGCTGCGCTGACCCCGAGCGGATGGCCGAG 197  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 198 GGTGGCTTCATCCACTGCCCCCACTGAGAACGACGACCTGGCCCACTGTTCTTCTG 257  
QY 61 PheYsGluLeuGluGlyTirpGluProAspAspPhePheIleGluGluHis 80  
DB 258 TTCAAGAGAGCTGAGGCTGGAGCCAGATGACGCCCAAGAGGAACATAAAGACAT 317  
QY 81 SerSerGlyCysAlaPheLeuSerValIleYsGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 318 TCGTCCGGTGGCTTCTCTTCTCTCTCAAGAAAGCATTTGAAGATTAACTTGGTGA 377  
QY 101 PheLeuYsLeuAspArgGluArgAlaYsAsnYsIleAlaYsGluThrAsnAsnYs 120  
DB 378 TTTTGAAGCTGAGCAGAGAAAGACCAAAATTTGCAAGGAACCAACATTAAG 437  
QY 121 LysYsGluPheGluGluThrAlaYsYsValaArgArgAlaIleGluGluLeuAlaAla 140  
DB 438 AAGAAAGATTGAGGAAACTGCGAAGAAAGTGGCGCTGCCATCGAGACGCTGGCTCC 497

QY 141 MetAsp 142  
DB 498 ATGCAT 503  
RESULT 14  
AM409816  
LOCUS  
DEFINITION  
AM409816 613 bp mRNA linear EST 29-JUN-2000  
fb02e10.x1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:296115 5',  
mRNA sequence.  
AM409816  
AM409816  
VERSION  
AM409816.1 GI:6935357  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
NIH-MGC http://mhc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ULNL at:  
www.bio.lnh.gov/bbrp/image/image.html  
Plate: LNCM55 row: I column: 20  
Seq primer: -21M3 forward primer (ABI).

## FEATURES

## source

1. 613  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:296115"  
/tissue\_type="Thadomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 17"  
/note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;  
Site 2: XhoI; cDNA made by oligo-dt priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

Alignment Scores:  
Pred. No.: 9, 216-149 Length: 613  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-690-825-34 (1-142) x AM409816 (1-613)

QY 1 MetGlyAlaProThrLeuProProAlaTirpGlnProPheLeuYAspHisArgIleSer 20  
DB 53 ATGGGTGCCCCGAGCTTCCCTGCGAGCCCTTTCAGAGCAACCGCATCTCT 112  
QY 21 ThrPheYAsnTirpProPheLeuGluGlyCysAlaCysTirpProGluArgMetAlaGlu 40  
DB 113 ACATTCAAGAACTGGCCCTCTTGTGAGGGCTGCGCTGACCCCGAGCGGATGGCCGAG 172  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 173 GGTGGCTTCATCCACTGCCCCCACTGAGAACGACGACCTGGCCCACTGTTCTTCTG 232



QY 61 PhelysGluLeuGluGlyTTPGluProAspAspProIleGluGluHislySlySHis 80  
 Db 233 TTCAGAGAGCTGGAGAGCTGGAGCCAGATAGCAGCCCATAGAGCAATTAAGGCAAT 292  
 QY 81 SerSerGlyCysAlaPheLeuSerValIlySlyGlnPheGluGluLeuThrIleGlyGlu 100  
 Db 293 TCGCCCGCTGCGCTTCTTCTTCTGTCAGAGAGATTGAAGATTTAACCTTGTGA 352  
 QY 101 PheLeuIlyLeuAspArgIlyArgAlaIlyAsnIlyIleAlaIlyGluThrAsnAsnIly 120  
 Db 353 TTTTGAACCTGACAGAGAGAGCCAGACCAATTTGCAAGAGAACCAATTAAG 412  
 QY 121 LysIlyGluPheGluGluGluThrAlaIlyIlyValArgArgAlaIleGluGluIleAlaAla 140  
 Db 413 AAGAAAGAAATTGTGAAGAACTGCGAAGAAAGTGGCGCTGCATCGAGCAGCTGGCTGCC 472  
 QY 141 MetAsp 142  
 Db 473 ATGGAT 478

RESULT 15  
 CBI35918 624 bp mRNA linear EST 29-JAN-2003  
 LOCUS K-EST0188232 LSHLK1 Homo sapiens cDNA clone LSHLK1-27-C08 5', mRNA  
 DEFINITION sequence.  
 CBI35918  
 VERSION CBI35918.1 GI:28103005  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 624)  
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.W., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 21C Frontier Korean EST Project 2001  
 Unpublished (2002)  
 Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of BioScience & Biotechnology  
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsun@mail.krdb.re.kr  
 Plate: 27 row: C column: 08  
 High quality sequence stop: 624.  
 Location/Qualifiers

FEATURES  
 SOURCE  
 1..624  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="LSHLK1-27-C08"  
 /sex="M"  
 /cell\_line="HK-1"  
 /lab\_host="TOP10"  
 /clone\_id="LSHLK1"  
 /note="Organ: Liver; Vector: pCNS-D2; Site: 1: EcoRI;  
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then deacapped  
 with tobacco acid pyrophosphatase (TAP). The deacapped  
 intact mRNA was ligated with DNA-RNA linker including  
 EcoRI site by treatment of T4 RNA ligase and the first  
 strand cDNA was synthesized from oligo dt-selected mRNA by  
 priming with dt-tailed vector. The dt-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was  
 circularized with E. coli DNA ligase after digestion of  
 EcoRI which site is also included in vector. An RNA strand  
 converted to a DNA strand by Okayama-Berg method. The  
 obtained cDNA vectors were used for transformation of  
 competent cells E. coli TOP10 by electroporation method.  
 The cDNA libraries constructed by this method are

ORIGIN full-length enriched cDNA library."

Alignment Scores:  
 Pred. No.: 9,356-149 Length: 624  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 14 Gaps: 0

US-09-690-825-34 (1-142) x CBI35918 (1-624)

QY 1 MetGlyAlaProThrLeuProProAlaThrGlnProPheLeuIlyAspHisArgIleSer 20  
 Db 25 ATGGGTGCCGAGAGCTGGCCCTGCTGGAGCCCTTTCTCAAGACACCGCATCTCT 84  
 QY 21 ThrPheIlyAsnIlyProPheLeuGluGlyCysAlaCysIlyThrProGluArgMetAlaGlu 40  
 Db 85 ACATTCAGAACTGGCCCTTCTTGAAGGCTGGCCCTGCACCCCGAGCGATGGCCGAG 144  
 QY 41 AlaGlyPheIleHisCysPProThrGluGlnGluProAspIleAlaGluCysPheCys 60  
 Db 145 GGTGCTTCATCTCACTGCCCCCTAGAACGAGCCAGCTGGCCAGTGTCTTCTGCG 204  
 QY 61 PheIlyGluLeuGluGlyTTPGluProAspAspProIleGluGluHislySlySHis 80  
 Db 205 TTCAGAGAGCTGGAGAGCTGGAGCCAGATAGACGCCCATAGAGAACTAAAGCAT 264  
 QY 81 SerSerGlyCysAlaPheLeuSerValIlySlyGlnPheGluGluLeuThrIleGlyGlu 100  
 Db 265 TCGTCCGCTGCGCTTCTTCTTCTGTCAGAGAGCTTGAAGATTAACCTTGTGA 324  
 QY 101 PheLeuIlyLeuAspArgIlyArgAlaIlyAsnIlyIleAlaIlyGluThrAsnAsnIly 120  
 Db 325 TTTTGAACCTGACAGAGAGAGCCAGACCAATTTGCAAGAGAACCAATTAAG 384  
 QY 121 LysIlyGluPheGluGluGluThrAlaIlyIlyValArgArgAlaIleGluGluIleAlaAla 140  
 Db 385 AAGAAAGAAATTGTGAAGAACTGCGAAGAAAGTGGCGCTGCATCGAGCAGCTGGCTGCC 444  
 QY 141 MetAsp 142  
 Db 445 ATGGAT 450

RESULT 16  
 BE292733 629 bp mRNA linear EST 13-JUL-2000  
 LOCUS BE292733 601105764P1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:2988543 5',  
 DEFINITION mRNA sequence.  
 BE292733  
 VERSION BE292733.1 GI:9175339  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 629)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA library Preparation: Ling Hong/Rubin Laboratory  
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: NIH Intramural Sequencing Center  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNL at: image.llnl.gov  
 Plate: LINC77 row: P column: 16  
 High quality sequence stop: 629.  
 Location/Qualifiers

FEATURES  
 source  
 1..629

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2988543"
/tissue_type="adenocarcinoma cell line"
/lab_host="MDH108 (phage-resistant)"
/clone_lib="NIM_MGC_15"
/notes="Organ: colon; Vector: pOT87; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 9,41e-149 Length: 629
Score: 142.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

```

US-09-690-825-34 (1-142) x BE292733 (1-629)

```

QY 1 MetGlyAlaProThleuProProAlaTrpGlnProPheLeuYsaPHisArgIleSer 20
DB 6 ATGGGTGGCCGACGCTTCCCTCCGCGAGCCCTTTCAGAGACCAACCGCATCTCT 65
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 66 ACATTCAAGAACTGGCCCTTCTTGAGGGCGCTGCGCTCCACCCGGAGCGGATGCCGAG 125
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB 126 GGTGGCTTCATCTCACTGCCCACTGAGAACAGCCAGACTGGCCCAAGTGTCTTCTTGC 185
QY 61 PheLysGluLeuGluGlyTTrpGluProAspAspAspProIleGluGluHisLysLysHis 80
DB 186 TTCAGAGAGCTGGAGAGGCTGGAGCCCAATGACACCCCATAGAGAACATMAAACAT 245
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB 246 TCGTCCGGTGGCTTCTTCTTCTGTCAGAGACGATTGAAGATTAAACCTTGTGAGAA 305
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAspAspLys 120
DB 306 TTTTGAAGACTGGACAGAGAAAGCCAGAAACAAATTGCAAGAGAAACCAACATPAG 365
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140
DB 366 AAGAAAGATTGAGAGAAACGCGAGAAAGAGGCGCGCTGCCATCGAGACGCTGCTGCC 425
QY 141 MetAsp 142
DB 426 ATGAT 431

```

## RESULT 17

```

BM788077 644 bp mRNA linear EST 05-MAR-2002
LOCUS BM788077
DEFINITION K-BST0067173 S10SNUI Homo sapiens cDNA clone S10SNUI-21-E12 5',
mRNA sequence.
ACCESSION BM788077
VERSION BM788077.1 GI:19136309
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 644)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

```

```

TITLE Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
JOURNAL Kim,Y.S.
Unpublished (2002)
CONTACT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoan-dong, Yuseong-gu, Daejeon 305-353, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 21 row: E column: 12
High quality sequence stop: 644.
Location/Qualifiers
1..644
source

```

## FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S10SNUI-21-E12"
/sex="M"
/tissue_type="Stomach"
/cell_type="Lymphoblast-like"
/cell_line="SNU-1"
/lab_host="SNU-1"
/clone_lib="S10SNUI"
/notes="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
Site 2: XhoI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deapped
with tobacco acid pyrophosphatase (TAP) and the deapped
intact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dt primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfiI and
cloned into DraIII- digested pME18-FL3 vector. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 9,61e-149 Length: 644
Score: 142.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

```

US-09-690-825-34 (1-142) x BM788077 (1-644)

```

QY 1 MetGlyAlaProThleuProProAlaTrpGlnProPheLeuYsaPHisArgIleSer 20
DB 45 ATGGGTGGCCGACGCTTCCCTCCGCGAGCCCTTTCAGAGACCAACCGCATCTCT 104
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 105 ACATTCAAGAACTGGCCCTTCTTGAGGGCGCTGCGCTCCACCCCGAGCGGATGCCGAG 164
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB 165 GGTGGCTTCATCTCACTGCCCACTGAGAACAGCCAGACTGGCCCAAGTGTCTTCTGCC 224
QY 61 PheLysGluLeuGluGlyTTrpGluProAspAspAspProIleGluGluHisLysLysHis 80
DB 225 TTCAGAGAGCTGGAGAGGCTGGAGACCAATGACGCCCAATAGAGAAACAAAAGCAT 284
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB 285 TCGTCCGGTGGCTTCTTCTTCTGTCAGAGACGATTGAAGATTAAACCTTGTGAGAA 344
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAspAspLys 120

```

```

Db      345 TTTTGAACCTGCAGAGAAAGCCAGACAAATTTGCAAGAAACCAATTAAG 404
Qy      121 LysylsGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAla 140
Db      405 AAGAAAGATTGTGAGAACTGCAGAGAAAGTGCCTGTCATCGACAGCTGGTGGC 464
Qy      141 MetAsp 142
Db      465 ATGGAT 470

RESULT 18
B1252508      646 bp      mRNA      linear      EST 17-JUN-2001
LOCUS      602952974.F1 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:5087178 5',
DEFINITION      mRNA sequence.
ACCESSION      B1252508
VERSION      B1252508.1 GI:14803034
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 646)
AUTHORS      NIH-MGC http://mgs.nci.nih.gov/.
TITLES      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: CGAP (Stanford)
              CDNA Library Preparation: Ling Hong/Rubin Laboratory
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: LINC844 row: c column: 19
              High quality sequence stop: 644.
              Location/Qualifiers
                1..646
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5087178"
                /tissue_type="hepatocellular carcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_1lb="NIH MGC 100"
                /note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
                EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                into EcoRI/XhoI sites using the following 5' adaptor:
                GGACGAG(G). Size-selected >500bp for average insert size
                1.8kb. Library constructed by Ling Hong in the laboratory
                of Gerald M. Rubin (University of California, Berkeley)
                using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                II RT (Life Technologies). Note: this is a NIH_MGC
                Library."

```

```

Db      116 ACATTCAAACTGGCCCTCTTGTAGAGGCTCGCTGCACCCCGAGGGATGGCCAG 175
Qy      41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60
Db      176 GCTGCTTATCATCATCTGCCCTCCACTGAGACGAGCCAGACTTGGCCAGTCTTCTTCC 235
Qy      61 PheLysGluLeuGluGlyTyrGluProAspAspProIleGluGlnHisLysLysHis 80
Db      236 TTCAAGAGCTGGAAGGCTGGAGCCAGATGACGACCCCATAGAGAAATATAAAGCAT 295
Qy      81 SerSerGlyCysAlaPheLeuSerValLysLysGluPheGluGlnLeuThrLeuGlyGlu 100
Db      296 TCGTCCGAGTGGCGCTTCTTCTTGTCAAGAAAGCTTTGAAGAAATTAACCTTGCTGA 355
Qy      101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
Db      356 TTTTGAACCTGCAGAGAAAGCCCAAGAACTTGCAGAAAGAAACCAATTAAG 415
Qy      121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAla 140
Db      416 AAGAAAGATTGTGAGAAAGCTGCAGAAAGATGCGCGCTGTCATGACAGCTGGCTGCC 475
Qy      141 MetAsp 142
Db      476 ATGGAT 481

RESULT 19
B2796084      682 bp      mRNA      linear      EST 20-SEP-2000
LOCUS      601591091.F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3945214 5',
DEFINITION      mRNA sequence.
ACCESSION      B2796084
VERSION      B2796084.1 GI:10217282
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 682)
AUTHORS      NIH-MGC http://mgs.nci.nih.gov/.
TITLES      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: DCTD/DRP
              CDNA Library Preparation: Ling Hong/Rubin Laboratory
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
              Plate: LINC803 row: e column: 23
              High quality sequence start: 3
              High quality sequence stop: 681.
              Location/Qualifiers
                1..682
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:3945214"
                /tissue_type="small cell carcinoma"
                /cell_line="MGC3"
                /lab_host="DH10B (phage-resistant)"
                /clone_1lb="NIH MGC 7"
                /note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
                EcoRI; cDNA made by oligo-dT priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GGACGAG(G). Size-selected >500bp for average
                insert size 1.8kb. Library constructed by Ling Hong in
                the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies)."
```

## ORIGIN

## Alignment Scores:

```

Pred. No.:      9.63e-149      Length:      646
Score:          142.00      Matches:      142
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:            12      Gaps:      0

```

US-09-690-825-34 (1-142) x B1252508 (1-646)

```

Qy      1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db      56 ATGGATCCCCCGACGTTGCCCTTGGCTGAGAGCCCTTTCAGAGACCCGACATCTCT 115
Qy      21 ThrPheLysAsnTyrProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGln 40

```

## ORIGIN

## Alignment Scores:

Pred. No.:	1,036-148	Length:	698
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-690-825-34 (1-142) x BE796084 (1-682)

```
OY 1 MetGlyAlaProThrLeuProProAlaTTPGlnProPheLeuLysAspHisArgIleSer 20
DB 23 ATGGGTGCCCCGACGCTTCCCTGCTGCGAGCCCTTCTCAAGAGCCACCCGACTCTCT 82
OY 21 ThrPheLysAsnTTPProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 83 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGCCCTGCACCCCGGAGCGGATGGCCGAG 142
OY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuIleGluCysPhePheCys 60
DB 143 GCTGGCTTCATCCACTGCCCCCATGAGACGAGCCAGACTTGGCCCGGCTTCTTCTGTC 202
OY 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluGluHis 80
DB 203 TTCAAGAGAGCTGGAAGCTGGAGCGAGATGACGCCATAGAGAACTAAAGAGCAT 262
OY 81 SerSerGlyCysAlaPheLeuSerValIleLysGlnPheGluGluLeuThrLeuGlu 100
DB 263 TCGTCGGGTGGCGCTTCTTCTTCTGTCAGAGAGAGATTGAACCTTGGTGA 322
OY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
DB 323 TTTTGAACCTGGACAGAGAAAGAGCCAGAAACAAATTGCAAGAGAAACCAACATAG 382
OY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140
DB 383 AAGAAAGAAATTGAGAGAACTGCGAAGAAAGTGGCGCGTCCATGAGCAGCTGGCTCC 442
OY 141 MetAsp 142
DB 443 ATGGAT 448
```

RESULT 20  
BE796061 698 bp mRNA linear EST 20-SEP-2000  
LOCUS 601591059F1 NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:3945206 5',  
DEFINITION mRNA sequence.

ACCESSION BE796061  
VERSION BE796061.1 GI:10217259

KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 698)

REFERENCE NIH-MGC http://mgi.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cga@bbs-remail.nih.gov  
Tissue Procurement: DCM/DTP

CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at: image.jnl.gov  
plate: LNC803 row: e column: 15

High quality sequence start: 2  
High quality sequence stop: 698.

FEATURES  
1..698  
Location/Qualifiers

source  
/organism="Homo sapiens"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,036-148 Length: 698  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-690-825-34 (1-142) x BE796061 (1-698)

```
OY 1 MetGlyAlaProThrLeuProProAlaTTPGlnProPheLeuLysAspHisArgIleSer 20
DB 22 ATGGGTGCCCCGACGCTTCCCTGCTGCGAGCCCTTCTCAAGAGCCACCCGACTCTCT 81
OY 21 ThrPheLysAsnTTPProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 82 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGCCCTGCACCCCGGAGCGGATGGCCGAG 141
OY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuIleGluCysPhePheCys 60
DB 142 GCTGGCTTCATCCACTGCCCCCATGAGACGAGCCAGCTTGGCCCGGCTTCTTCTGTC 201
OY 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluGluHis 80
DB 202 TTCAAGAGAGCTGGAAGCTGGAGCGAGATGACGCCATAGAGAACTAAAGAGCAT 261
OY 81 SerSerGlyCysAlaPheLeuSerValIleLysGlnPheGluGluLeuThrLeuGlu 100
DB 262 TCGTCGGGTGGCGCTTCTTCTTCTGTCAGAGAGAGCTTGAAGATTGAACCTTGGTGA 321
OY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
DB 322 TTTTGAACCTGGACAGAGAAAGAGCCAGAAACAAATTGCAAGAGAAACCAACATAG 381
OY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140
DB 382 AAGAAAGAAATTGAGAGAACTGCGAAGAAAGTGGCGCGTCCATGAGCAGCTGGCTGCC 441
OY 141 MetAsp 142
DB 442 ATGGAT 447
```

RESULT 21  
BF219993 703 bp mRNA linear EST 08-NOV-2000  
LOCUS 601296980F1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:2964713 5',  
DEFINITION mRNA sequence.

ACCESSION BF219993  
VERSION BF219993.1 GI:11126087

KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 703)  
NIH-MGC http://mgi.nci.nih.gov/.

Tue Aug 17 05:54:54 2004

us-09-690-825-34.oligo.rst

Page 15

TITLE  
JOURNAL  
COMMENT  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: IRAL1 row: p column: 19  
High quality sequence stop: 656.

FEATURES  
source  
1. 703  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2964713"  
/tissue\_type="rhabdomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 17"  
/note="Organ: muscle, Vector: pOTB7, Site\_1: EcoRI;  
Site\_2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.04e-148 Length: 703  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-690-825-34 (1-142) x BP219993 (1-703)  
QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 16 ATGGGTGCCCCGAGAGTTCGCCCTGGAGAGCCCTTCTCAAGACCAACCGCATCTT 75  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 76 ACATTCAAGAACTGACCTCTTCTTGAAGGCTGCGCTGACCCCGAGCGGATGCGCGAG 135  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 136 GCTGGCTTCATCCACTGCCCCACGAGACGACCAACTGGCCCCAGGTTCTTCTCTC 195

QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluLysIleLysLysHis 80  
DB 196 TTCAAGAGCTGGAAGGCTGGAGCCAGATGACACCCCATAGAGAACTAAAGAGCAT 255  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGluGlu 100  
DB 256 TCGTCGGGTTCGCTTCTTCTTCTGTCAGAGAGAGTTGAAGATTAAACCTGGTGA 315  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
DB 316 TTTTGAAGCTGACGAGAAAGCCAGAAACAAATTGCAAGAGAAACCAACATTAAG 375  
QY 121 LysLysGluPheGluGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAla 140  
DB 376 AAGAAAGATTTTGAAGAACTGCGAAGAAAGTGGCGCTGCATCGAGCGAGCTGCGC 435  
QY 141 MetAsp 142  
DB 436 ATGGAT 441  
RESULT 22

BG115707  
LOCUS 707 bp mRNA linear EST 30-JAN-2001  
DEFINITION 602317179F1 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:4417722 5',  
mRNA sequence.  
ACCESSION BG115707  
VERSION BG115707.1 GI:12609213  
KEYWORDS EST.  
ORGANISM Homo sapiens (human)

REFERENCE  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ARCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LHAM10150 row: m column: 19  
High quality sequence stop: 689.

FEATURES  
source  
1. 707  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4417722"  
/tissue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 88"  
/note="Organ: small intestine; Vector: pCMV-SPORT6;  
Site\_1: NotI, Site\_2: SalI; Cloned unidirectionally;  
oligo-dT primed. Average insert size 1.767 kb. Library  
enriched for full-length clones and constructed by Life  
Technologies. Note: this is a NIH\_MGC Library."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.04e-148 Length: 707  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-690-825-34 (1-142) x BG115707 (1-707)  
QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 27 ATGGGTGCCCCGAGAGTTCGCCCTGGAGAGCCCTTCTCAAGACCAACCGCATCTT 86  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 87 ACATTCAAGAACTGACCTCTTCTTGAAGGCTGCGCTGACCCCGAGCGGATGCGCGAG 146

QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 147 GCTGGCTTCATCCACTGCCCCACGAGACGACCAACTGGCCCCAGGTTCTTCTGC 206  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluLysIleLysLysHis 80  
DB 207 TTCAAGAGCTGGAAGGCTGGAGCCAGATGACACCCCATAGAGAACTAAAGAGCAT 266  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGluGlu 100  
DB 267 TCGTCGGGTTCGCTTCTTCTTCTGTCAGAGAGAGTTGAAGATTAAACCTGGTGA 326  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

```

FEATURES
  Faraday Avenue Genoscope sequence ID : CSUD0006BR02021.
  Location/Qualifiers
    1..726
      /organism="Homo sapiens"
      /mol_type="rRNA"
      /toll_type="rRNA"

```

```

/clone="CS0DH006Y104"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"

```

/note=Vector: pcwvSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pcwvSPORT 6 vector. Library was not normalized.

## ORIGIN

## Alignment Scores:

Pred. No.:	1.08e-148	Length:	726
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-09-690-825-34 (1-142) x BX421971 (1-726)

```

QY      1 MetGlyAlaProThrLeuProProAlaATrPGlnProPheLeuLysAspHisArgIleSer 20
DB      127 ATGGGTGCCCCGACGCTTGCCTCCCTGCGACGCTTTCTCAAGACACCGCATCTCT 186
QY      21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB      187 ACATTCAAGAACTGGCCCTTCTTGAGAGGCTGCGCTGCACCCGGAGCGGATGCGCGAG 246
QY      41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB      247 GCTGGCTTCATCCACTGCCCTGACAGACGACGACGCTGCGCCAGTGTCTTCTTCTGC 306
QY      61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80
DB      307 TTCAAGAGGCTGGAAGGCTGGAGCCAGATGACGACCCCATAGAGACATTAAGCAT 366
QY      81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB      367 TCGTCCGTTGGCGCTTCTTCTTCTCAAGAAAGCGTTGAAGAAATTAACCTTGCTGAA 426
QY      101 PheLeuLysLeuAspArgGluArgAlaLysLysLysLysLysLysLysLysLysLys 120
DB      427 TTTTGAACACTGAGAGAAAGACCAAGAAATTCGAAAGAAACCAACCAATTAAG 486
QY      121 LysLysGluPheGluGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAla 140
DB      487 AAGAAAGAAATTTGAGAGAACTGCCGAAAGAAAGTGGCCCTGCATGACCACTGCTGCC 546
QY      141 MetAsp 142
DB      547 ATGGAT 552

RESULT 25
BE907403      742 bp      mRNA      linear      EST 20-OCT-2000
LOCUS      601500136F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3901880 5',
DEFINITION      mRNA sequence.
ACCESSION      BE907403
VERSION      BE907403.1 GI:10400927
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 742)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLES      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:

```

http://image.llnl.gov  
 Plate: LL98703 row: h column: 09  
 High quality sequence stop: 707.  
 Location/Qualifiers  
 1..742  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3901880"  
 /tissue\_type="epithelioid carcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_70"  
 /note="Organ: pancreas; Vector: pcwv-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.1 kb. Library constructed by Life  
 Technologies."

## FEATURES

## Source

## ORIGIN

## Alignment Scores:

Pred. No.:	1.08e-148	Length:	742
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-690-825-34 (1-142) x BE907403 (1-742)

```

QY      1 MetGlyAlaProThrLeuProProAlaATrPGlnProPheLeuLysAspHisArgIleSer 20
DB      30 ATGGGTGCCCCGACGCTTGCCTCCCTGCGACGCTTTCTCAAGACACCGCATCTCT 89
QY      21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB      90 ACATTCAAGAACTGGCCCTTCTTGAGAGGCTGCGCTGCACCCGGAGCGGATGCGCGAG 149
QY      41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB      150 GCTGGCTTCATCCACTGCCCTGACAGACGACGACGCTGCGCCAGTGTCTTCTTCTGC 209
QY      101 PheLeuLysLeuAspArgGluArgAlaLysLysLysLysLysLysLysLysLysLys 80
DB      210 TTCAAGAGGCTGGAAGGCTGGAGCCAGATGACGACCCCATAGAGAAATTAAGCAT 269
QY      61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80
DB      270 TCGTCCGTTGGCGCTTCTTCTTCTGTCAGAAAGCAATTCGAAATTAACCTTGCTGAA 329
QY      81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB      270 TCGTCCGTTGGCGCTTCTTCTTCTGTCAGAAAGCAATTCGAAATTAACCTTGCTGAA 329
QY      101 PheLeuLysLeuAspArgGluArgAlaLysLysLysLysLysLysLysLysLysLys 120
DB      330 TTTTGAACACTGAGAGAAAGACCAAGAAATTCGAAAGAAACCAACCAATTAAG 389
QY      121 LysLysGluPheGluGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAla 140
DB      390 AAGAAAGAAATTTGAGAGAACTGCCGAAAGAAAGTGGCCCTGCATGACGAGCTGCTGCC 449
QY      141 MetAsp 142
DB      450 ATGGAT 455

RESULT 26
BM020628      760 bp      mRNA      linear      EST 30-OCT-2001
LOCUS      603643029F1 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5419087 5',
DEFINITION      mRNA sequence.
ACCESSION      BM020628
VERSION      BM020628.1 GI:16534982
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 760)

```

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LICM1871 row: 1 column: 08  
 High quality sequence stop: 722.

## FEATURES

Location/Qualifiers

1..760  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5419087"  
 /tissue\_type="astrocytoma grade IV, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1ib="NIH MGC 98"  
 /note="Organ: brain; Vector: pOTB7, Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,11e-148 Length: 760  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-690-825-34 (1-142) x BM020628 (1-760)

QY 1 MetG|A|A|P|ro|th|re|u|p|ro|f|o|a|t|p|g|in|p|ro|p|he|u|le|u|s|a|s|p|h|s|a|r|g|l|e|s|e|r 20  
 Db 33 A|T|G|G|G|G|C|C|C|G|A|G|T|G|C|C|C|C|T|G|C|C|G|A|G|C|C|T|T|T|C|T|A|A|G|A|C|C|A|C|C|G|C|A|T|C|T 92  
 QY 21 T|h|P|h|e|l|y|s|a|s|n|t|p|p|ro|p|he|u|l|e|u|g|l|y|C|y|s|a|l|A|C|y|s|T|h|P|ro|g|l|u|a|r|g|e|t|a|g|u 40  
 Db 93 A|C|A|T|T|C|A|A|G|A|C|T|G|G|C|C|T|T|C|T|T|G|A|G|G|G|C|T|G|C|A|C|C|C|G|A|G|G|A|T|G|G|C|C|G|A 152  
 QY 41 A|l|A|G|l|P|h|e|l|l|e|h|s|C|y|s|P|ro|t|h|r|g|l|u|a|e|n|g|l|u|p|ro|A|s|p|l|e|u|A|l|A|G|l|n|C|y|s|P|h|e|p|h|e|C|y|s 60  
 Db 153 G|C|T|G|G|T|T|C|A|T|C|A|C|A|C|G|C|C|C|C|A|T|G|A|A|C|G|A|C|G|A|C|T|T|G|G|C|C|A|G|T|T|C|T|T|G|C 212  
 QY 61 P|h|e|l|y|s|e|l|u|e|u|g|l|y|T|P|G|l|u|P|ro|A|s|p|A|s|P|ro|l|e|g|l|u|h|s|l|y|s|l|y|s|H|s 80  
 Db 213 T|T|C|A|A|G|A|G|C|T|G|A|A|G|C|T|G|G|A|G|C|A|T|G|A|G|A|C|C|C|A|T|G|A|G|A|C|A|T|A|A|A|A|G|C|A|T 272  
 QY 81 S|e|S|e|G|l|y|C|y|s|A|l|P|h|e|u|s|e|r|A|l|y|s|l|y|s|g|l|n|P|h|e|l|l|e|u|l|e|u|t|h|e|u|l|e|u|l|y|g|l|u 100  
 Db 273 T|C|G|T|C|G|G|T|G|G|C|C|T|T|C|T|T|C|T|G|T|C|A|A|G|A|G|C|A|G|T|T|G|A|A|G|A|T|T|A|A|C|C|T|T|G|T|G|A 332  
 QY 101 P|h|e|l|l|y|s|e|u|s|a|s|p|r|g|l|u|a|r|g|a|l|y|s|a|s|u|l|y|s|l|e|a|l|y|s|g|l|n|u|t|h|s|a|s|u|l|y|s 120  
 Db 333 T|T|T|T|G|A|A|C|T|G|A|C|A|G|A|A|G|C|A|A|G|C|A|A|A|T|T|G|C|A|A|G|A|A|C|C|A|C|A|T|A|A|G 392  
 QY 121 L|y|L|y|e|g|l|u|P|h|e|l|l|u|t|h|a|r|a|l|y|s|l|y|s|V|a|l|a|r|g|a|r|a|l|l|e|g|l|u|l|e|u|A|a|l|a 140  
 Db 393 A|A|G|A|A|G|A|T|T|C|A|G|A|A|A|C|T|G|C|A|A|A|G|T|G|C|G|C|G|C|A|T|G|A|G|A|G|T|G|C|T|G|C 452  
 QY 141 M|e|t|A|s|p 142

Db 453 ATGAT 458

RESULT 27  
 BG682887 766 bp mRNA linear EST 01-MAY-2001  
 LOCUS 602651082F1 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:4761932 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG682887  
 VERSION BG682887.1 GI:13914271  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 766)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LICM1615 row: c column: 21  
 High quality sequence stop: 763.

## FEATURES

Location/Qualifiers

1..766  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4761932"  
 /tissue\_type="neuroblastoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1ib="NIH MGC 47"  
 /note="Organ: brain; Vector: pOTB7, Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,11e-148 Length: 766  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-690-825-34 (1-142) x BG682887 (1-766)

QY 1 MetG|A|A|P|ro|th|re|u|p|ro|f|o|a|t|p|g|in|p|ro|p|he|u|le|u|s|a|s|p|h|s|a|r|g|l|e|s|e|r 20  
 Db 56 A|T|G|G|T|C|C|C|G|A|G|T|G|C|C|C|C|T|G|C|C|G|A|G|C|C|T|T|T|C|T|A|A|G|A|C|C|A|C|C|G|C|A|T|C|T 115  
 QY 21 T|h|P|h|e|l|y|s|a|s|n|t|p|p|ro|p|he|u|l|e|u|g|l|y|C|y|s|a|l|A|C|y|s|T|h|P|ro|g|l|u|a|r|g|e|t|a|g|u 40  
 Db 116 A|C|A|T|T|C|A|A|G|A|C|T|G|G|C|C|T|T|G|A|G|G|C|T|G|C|A|C|C|C|G|A|G|G|A|T|G|G|C|C|G|A 175  
 QY 41 A|l|A|G|l|P|h|e|l|l|e|h|s|C|y|s|P|ro|t|h|r|g|l|u|a|e|n|g|l|u|p|ro|A|s|p|l|e|u|A|l|A|G|l|n|C|y|s|P|h|e|p|h|e|C|y|s 60  
 Db 176 G|C|T|G|G|T|T|C|A|T|C|A|C|A|C|G|C|C|C|C|A|T|G|A|A|C|G|A|C|G|A|C|T|T|G|G|C|C|A|G|T|T|C|T|T|G|C 235  
 QY 61 P|h|e|l|y|s|e|l|u|e|u|g|l|y|T|P|G|l|u|P|ro|A|s|p|A|s|P|ro|l|e|g|l|u|h|s|l|y|s|l|y|s|H|s 80



Db 236 TTCAAGAGCTGGAAGCTGGAGCCAGATGACGACCCCATAGAGAAACATMAAACAT 295

Qy 81 SerSerGIYCySAIAPheLeuSerValAlaLysLysGluPheGluGluLeuThrLeuGlyGlu 100

Db 286 TCGTCCGGTTCGGCTTTCTTCTGTCTCAAGAAAGCACTTTGAAGAAATTAAACCTTGATGA 355

Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120

Db 356 TTTTGAACCTGACAGAGAAAGAGCCCAAGAAATTCAGAAAGAAACCAACATTAAG 415

Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAla 140

Db 416 AAGAAAGAAATTGAGAAAGCTGCGAAGAAAGTCCCTGCATGAGACAGCTGCTGCC 475

Qy 141 MetAsp 142

Db 476 ATGGAT 481

RESULT 28

Bi771953 768 bp mRNA linear EST 25-SEP-2001

LOCUS 603058976F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5208259 5'

DEFINITION mRNA sequence.

ACCESSION Bi771953

VERSION Bi771953.1 GI:15763531

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 768)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Plate: L1AM1522 row: h column: 20

High quality sequence stop: 757.

Location/Qualifiers

1..768

FEATURES

source

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5208259"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_122"

/note="Organ: pooled lung and spleen; Vector: pCMV-SPORTS; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb. insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 1,12e-148 Length: 768

Score: 142.00 Matches: 142

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

US-09-690-825-34 (1-142) x Bi771953 (1-768)

Qy 1 MetGlyAlaProThrLeuProPheAlaATPGLNPProPheLeuLysAspHisArgIleSer 20

Db 48 ATGGAGGCCCCCAGCAGCTGGCCCCCTGCTGCGAGCCCTTCTCAAGAGCAACCGCATCTCT 107

Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyCySAIaCysThrProGluArgMetAlaGlu 40

Db 108 ACATTCAAGAAAGCTGCGCTTCTTGGAGGGCTCGCTGACCCCGAGGCGGATGGCCGAG 167

Qy 41 AlaGlyPheIleHisCysProThrArgLysGluAsnGluProAspLeuAlaGlnCysPhePheCys 60

Db 168 GCTGGCTTCATCCACAGCTGCCCCACAGAGAGCCAGACTTGAGCCCAAGTTCCTTCTG 227

Qy 61 PheLysGluPheGluGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80

Db 228 TTCAAGAGCTGGAAGCTGGAGCCAGATGACGACCCCATAGAGGACATMAAACAT 287

Qy 81 SerSerGIYCySAIAPheLeuSerValAlaLysLysGluPheGluGluLeuThrLeuGlyGlu 100

Db 288 TCGTCCGGTTCGGCTTTCTTCTGTCTCAAGAAAGCACTTTGAAGAAATTAAACCTTGATGA 347

Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120

Db 348 TTTTGAACCTGACAGAGAAAGAGCCCAAGAAATTCAGAAAGAAACCAACATTAAG 407

Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAla 140

Db 408 AAGAAAGAAATTGAGAAAGCTGCGAAGAAAGTCCCTGCATGAGACAGCTGCTGCC 467

Qy 141 MetAsp 142

Db 468 ATGGAT 473

RESULT 29

BG750227 772 bp mRNA linear EST 15-MAY-2001

LOCUS 602709025F1 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:4845582 5'

DEFINITION mRNA sequence.

ACCESSION BG750227

VERSION BG750227.1 GI:14060880

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 772)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Plate: L1C1685 row: a column: 07

High quality sequence stop: 765.

Location/Qualifiers

1..772

FEATURES

source

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4845582"

/tissue\_type="normal pigmented retinal epithelium"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_43"

/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5'

## ORIGIN

adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library. |"

## Alignment Scores:

Pred. No.:	1,12e-148	Length:	772
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-09-690-825-34 (1-142) x BG750227 (1-772)

```
QY 1 MetGlyAlaProThrLeuProAlaATPGLnProPheLeuLysAspHisArgIleSer 20
Db 16 ATGGGAGCGCCGACGTTGCCCTGCTGCGAGCCCTTCTCAGACACCGCATCTCT 75
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 76 ACATTAAAGACTGGCCCTTCTTGGAGGCTGCGCTGCAACCCGAGCGATGCCAG 135
QY 41 AlaGlyPheIleHisCysPProThrGluAengluProAspLeuAlaGlnCysPhePheCys 60
Db 136 GCTGCTTCATCCTACATGCCCCCACTGAGAACGAGCAGACTTGCCAGATTTCTTCTG 195
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisIleLysLysHis 80
Db 196 TTCAAGAGCTGGAGAGCTGGAGCCGATGACGACCCCATGAGAGAACATPAAAAAGCAT 255
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100
Db 256 TCGTCCGTTGGCGCTTCTCTTCTGTCAGAAAGCATTTGAAGATTAACTTGTGTGA 315
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120
Db 316 TTTTGAACCTGACAGAGAAAGAGCAAAATTTGCAAGAAACCAACATATAG 375
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140
Db 376 AAGAAAGATTGAGGAAACTGGAAGAAAGTGCCTGCGCATGACAGCTGCTGCC 435
QY 141 MetAsp 142
Db 436 ATGGAT 441
```

RESULT 30  
BU600185 785 bp mRNA linear EST 20-SEP-2002  
LOCUS AGENCOURT\_8870047 NIH\_MGC\_142 Homo sapiens cDNA clone IMAGE:6463579  
DEFINITION 5', mRNA sequence.  
ACCESSION BU600185  
VERSION BU600185.1 GI:23251944  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 785)  
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: NCI  
cDNA Library Preparation: Michael Brownstein Laboratory  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov  
plate: LICM655 row: e column: 20  
High quality sequence stop: 550.  
Location/Qualifiers

## FEATURES

source

1. 785

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6463579"

/tissue\_type="mixed (pool of 40 RNAs)"

/lab\_host="DH10B (T1-phage-resistant)"

/clone\_lib="NIH\_MGC\_142"

/note="Vector: pDNR-LIB; Site 1: SfiI (ggccatcaggcc);

Site 2: SfiI (ggccgcctcgcc); Double-stranded cDNA was

prepared from a pool of 40 cell line polyA+ RNAs (bladder

- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -

4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,

kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,

ovary - 1.3%, pharynx - 2.5%, prostate - 4.3%, salivary

gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were

used in cloning as follows:

5'-AGCAGAGGATCAACGACGAGCGCCATTCAGCGCGG-3' and

5'-ATTCTAGAGCGGAGCGCGCCATG-AT(30)NN-3'. Full-length

enriched library was constructed using the Clontech

Creator SMART kit and size-selected to contain the &gt;0.5 kb

size fraction (other fractions present in NIH\_MGC\_141).

Library created in the laboratory of M. Brownstein (NIH).

Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.:	1,14e-148	Length:	785
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-09-690-825-34 (1-142) x BU600185 (1-785)

```
QY 1 MetGlyAlaProThrLeuProAlaATPGLnProPheLeuLysAspHisArgIleSer 20
Db 58 ATGGGAGCGCCGACGTTGCCCTGCTGCGAGCCCTTCTCAGACACCGCATCTCT 117
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 118 ACATTCAAGACTGGCCCTTCTTGGAGGCTGCGCTGCAACCCGAGCGATGCCAG 177
QY 41 AlaGlyPheIleHisCysPProThrGluAengluProAspLeuAlaGlnCysPhePheCys 60
Db 178 GCTGCTTCATCCTACATGCCCCCACTGAGAACGAGCAGACTTGCCAGATTTCTTCTG 237
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisIleLysLysHis 80
Db 238 TTCAAGAGCTGGAGAGCTGGAGCCGATGACGACCCCATGAGAGAACATPAAAAAGCAT 297
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100
Db 298 TCGTCCGTTGGCGCTTCTCTTCTGTCAGAAAGCATTTGAAGATTAACTTGTGTGA 357
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120
Db 358 TTTTGAACCTGACAGAGAAAGAGCAAAATTTGCAAGAAACCAACATATAG 417
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140
Db 418 AAGAAAGATTGAGGAAACTGGAAGAAAGTGCCTGCGCATGACAGCTGCTGCC 477
QY 141 MetAsp 142
Db 478 ATGGAT 483
```

RESULT 31

```

BG335442      815 bp      mRNA      linear      EST 27-FEB-2001
LOCUS         602403907F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4541737 5',
DEFINITION    mRNA sequence.
ACCESSION     BG335442
VERSION       BG335442.1 GI:13141880
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE     1 NIH-MGC http://mgc.nci.nih.gov/.
              NIH-MGC http://mgc.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cga@bbs-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Ling Hong/Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLCM1220 row: m column: 02
              High quality sequence stop: 708.
FEATURES
  source
    1..815
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:4541737"
    /tissue_type="choriocarcinoma"
    /lab_host="DH10B (phage-resistant)"
    /clone_1lb="NIH_MGC_21"
    /note="Organ: placenta; Vector: pOTB7, Site_1: XhoI;
    Site_2: EcoRI; cDNA made by oligo-dT priming.
    Directionally cloned into EcoRI/XhoI sites using the
    following 5' adaptor: GGCACGAG(G). Size-selected >500bp
    for average insert size 1.8kb. Library constructed by
    Ling Hong in the laboratory of Gerald M. Rubin (University
    of California, Berkeley) using ZAP-cDNA synthesis kit
    (Stratagene) and Superscript II RT (Life Technologies)."
```

```

Alignment Scores:
Pred. No.:      1,176-148      Length:      815
Score:          142.00         Matches:      142
Percent Similarity: 100.00%     Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match:    100.00%         Indels:      0
DB:             12             Gaps:        0

US-09-690-825-34 (1-142) x BG335442 (1-815)

ORIGIN
1 MetGtAlaProThrLeuProProAlaTTPGlnProPhoLeuLysAspHisArgIleSer 20
13 ATGGGGGCCCGGCGGCTTGGCTGCGCTTCTTCAAGACACCGCATCTCT 72
21 ThrPhelAsnTTPProPhoLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
73 ACATTGAAGAACTGGCTTCTTGGAGGCTGGCTGACCCCGAGCGAGTGGCCGAG 132
41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
133 GCGGGCTTCATCCACTGCCCACTGAGACGAGCCGAGCTTGCGCCAGGTTCTTCTGCG 192
61 PheLysGluLeuGluGlyTTPGluProAspAspAspProIleGluGlnHisLysLysHis 80
193 TTCAAGAGAGCTGGAGGCTGGAGCGAGTGAAGACCCCATAGAGAACTAAAGACAT 252
81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100

FEATURES
  source
    1..815
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:4849640"
    /tissue_type="natural killer cells, cell line"
    /lab_host="DH10B (phage-resistant)"
    /clone_1lb="NIH_MGC_106"
    /note="Organ: blood; Vector: pOTB7, Site_1: XhoI; Site_2:
    EcoRI; cDNA made by oligo-dT priming. Directionally cloned
    into EcoRI/XhoI sites using the following 5' adaptor:
    GGCACGAG(G). Library constructed by Ling Hong in the
    laboratory of Gerald M. Rubin (University of California,
    Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
    Superscript II RT (Life Technologies). Note: this is a
    NIH_MGC library."
```

QY 21 ThPhelysaantTppropheuenglucyvalaCysTh-ProgluThMetalaclu 40  
 Db 93 ACATTCAAGAACTGGCCCTTCTTGAGGGCTGGCTGACCCCGAGCGCATGGCCGAG 152  
 QY 41 AlAGlYpHeilIehisCysProthrgluasngluProaspLeuAglncyspHeCys 60  
 Db 153 GCTGGCTTCATCCACTGCCCCCTGAGAACGAGCCGACTTGCCCCAGTGTCTTCTGC 212  
 QY 61 PhelYsgLuleuenglucyTTPgluProaspPaspProilegluLuhisylsYshis 80  
 Db 213 TTCAGAGAGCTGAGAGGCTGGAGGAGATGAGACCCCATAGAGAACTATAAAGCAT 272  
 QY 81 SerSerglYcysAlaPheleuserValylsYsgLInPhegluLuhenthrgluYglu 100  
 Db 273 TCGTCCGCTTCCCTTCTTCTGTCAGAGCACTTGAGAAATTAACCTTGCTGAA 332  
 QY 101 PhelYsgLuleuaspArgluArgAlaYsaenlyleAlaYsgLuthrsnaenly 120  
 Db 333 TTTTGAACCTGACAGAGAAAGAGCCAGAACAAATTTGCAAGAGAAACCAACATAG 392  
 QY 121 LysLysgluPhegluLuthrAlaYsValaYsValaYsAlailegluLuleuAla 140  
 Db 393 AAGAAAGAAATTTGAGAAAGCTGGAGAAAGTGGCCGTCATGAGCAGCTGGCTGCC 452  
 QY 141 MetAsp 142  
 Db 453 ATGGAT 458

RESULT 33  
 BM542207 821 bp mRNA linear EST 20-FEB-2002  
 LOCUS AGENCOURT 6436722 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5539841  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BM542207  
 VERSION BM542207.1 GI:18771512  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: ATCC/DCMD/DTF  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing By: The I.M.A.G.E. Consortium (LNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LLNL2234 row: P column: 18  
 High quality sequence stop: 448.  
 Location/Qualifiers  
 1..821  
 source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5539841"  
 /rname="mRNA"  
 /rseq\_id="5539841"  
 /rsize\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC\_72"  
 /note="Organ: skin; Vector: PCMV-SF076; Site: 1; NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2 kb. Library constructed by Life  
 Technologies."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.18e-148 Length: 821

Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-09-690-825-34 (1-142) x BM542207 (1-821)

QY 1 MetGlyAlaProthrluProProAlaATPgluProPheleuYsaPhisArgIleSer 20  
 Db 8 ATGGGTGCCCGAGAGTGGCCCTGCTGCTGAGACCCCTTCTCAAGAGCACCCGACATCT 67  
 QY 21 ThPhelysaantTppropheuenglucyvalaCysTh-ProgluThMetalaclu 40  
 Db 68 ACATTCAAGAACTGGCCCTTCTTGAGGGCTGGCTGACCCCGAGCGCATGGCCGAG 127  
 QY 41 AlAGlYpHeilIehisCysProthrgluasngluProaspLeuAglncyspHeCys 60  
 Db 128 GCTGGCTTCATCCACTGCCCCCTGAGAACGAGCCGACTTGCCCCAGTGTCTTCTGC 187  
 QY 61 PhelYsgLuleuenglucyTTPgluProaspPaspProilegluLuhisylsYshis 80  
 Db 188 TTCAGAGAGCTGAGAGGCTGGAGGAGATGAGACCCCATAGAGAACTATAAAGCAT 247  
 QY 81 SerSerglYcysAlaPheleuserValylsYsgLInPhegluLuhenthrgluYglu 100  
 Db 248 TCGTCCGCTTCCCTTCTTCTGTCAGAGCACTTGAGAAATTAACCTTGCTGAA 307  
 QY 101 PhelYsgLuleuaspArgluArgAlaYsaenlyleAlaYsgLuthrsnaenly 120  
 Db 308 TTTTGAACCTGACAGAGAAAGAGCCAGAACAAATTTGCAAGAGAAACCAACATAG 367  
 QY 121 LysLysgluPhegluLuthrAlaYsValaYsValaYsAlailegluLuleuAla 140  
 Db 368 AAGAAAGAAATTTGAGAAAGCTGGAGAAAGTGGCCGTCATGAGCAGCTGGCTGCC 427  
 QY 141 MetAsp 142  
 Db 428 ATGGAT 433

RESULT 34  
 AUI30493 822 bp mRNA linear EST 01-ANG-2002  
 LOCUS AUI30493 NT2RP3 Homo sapiens cDNA clone NT2RP3000936 5', mRNA  
 DEFINITION sequence.  
 ACCESSION AUI30493  
 VERSION AUI30493.1 GI:10990847  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 822)  
 AUTHORS Oca,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,  
 Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and  
 Iisaga,T.  
 TITLE HRI human cDNA project  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Takao Iisaga  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genomics@hri.co.jp  
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
 Research Institute; cDNA library construction: Department of  
 Virology, Institute of Medical Science, University of Tokyo, and  
 Helix Research Institute.  
 Location/Qualifiers  
 1..822  
 source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"

/db\_xref="taxon:9606"  
/clone="NT2RP3000936"  
/cell\_type="retalocarcinoma"  
/cell\_line="NT2"  
/clone\_id="NT2RP3"  
/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor  
cells after 2-weeks retinoic acid (RA) induction"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,18e-148 Length: 822  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x AU130493 (1-822)

QY 1 MetGlyAlaProThrLeuProProAlaTTPGInProPheLeuLysAspHisArgIleSer 20  
DB 40 ATGGGTGCCCGAGCTTCCCTGCTGCGAGCCCTTTCTCAAGACACCGCATCTCT 99  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 100 ACATTCAAGAACTGGCCCTTCTTGAGAGGCTGCGCTGCACCCCGAGCGGATGCGCGAG 159  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 160 GCTGCTTCATCCATCCATGCCCCACTGAGAACAGACGAGCTGGCCAGTGTCTCTCTGC 219  
QY 61 PheLysGluLeuGluGlyTTPGInProAspAspProIleGluGluHisLysLysHis 80  
DB 220 TTCAAGAGAGCTGAGAGCTGGAGCCAGATGACGCCCATAGAGGAACATTAAGAAGAT 279  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 280 TCGTCCGGTGGCGCTTCTCTCTCTCTCAAGAAAGCAGTTGAAGATTAACTTGGTGA 339  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
DB 340 TTTTGAACCTGAGACAGAAAGAACCCAGAACAAATTTGCAAGAAACCAACATTAAG 399  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140  
DB 400 AAGAAAGAAATTGAGAAACTGCGAAGAAAGTGGCGCTGCATGAGCACTGGCTGCC 459  
QY 141 MetAsp 142  
DB 460 ATGGAT 465

RESULT 35  
LOCUS BU171608 826 bp mRNA linear EST 04-SEP-2002  
DEFINITION AGENCOURT\_7945706 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6149893  
5', mRNA sequence.  
ACCESSION BU171608.1 GI:22685592  
VERSION BU171608  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 826)  
NIH-MGC http://mgs.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMIL)  
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILMIL at:  
http://image.llnl.gov  
Plate: LLM13483 row: k column: 14  
High quality sequence stop: 733.  
Location/Qualifiers  
1. 826  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6149893"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH MGC 72"  
/note="Organ: skin; Vector: pCMV-Sport6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

## FEATURES

source

## ORIGIN

Alignment Scores:  
Pred. No.: 1,19e-148 Length: 826  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x BU171608 (1-826)

QY 1 MetGlyAlaProThrLeuProProAlaTTPGInProPheLeuLysAspHisArgIleSer 20  
DB 25 ATGGGTGCCCGAGCTTCCCTGCTGCGAGCCCTTTCTCAAGACACCGCATCTCT 84  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 85 ACATTCAAGAACTGGCCCTTCTTGAGAGGCTGCGCTGCACCCCGAGCGGATGCGCGAG 144  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 145 GCTGCTTCATCCATCCATGCCCCACTGAGAACAGACGAGCTGGCCAGTGTCTCTCTGC 204  
QY 61 PheLysGluLeuGluGlyTTPGInProAspAspProIleGluGluHisLysLysHis 80  
DB 205 TTCAAGAGAGCTGAGAGCTGGAGCCAGATGACGCCCATAGAGGAACATTAAGAAGAT 244  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 265 TCGTCCGGTGGCGCTTCTCTCTCTCTCAAGAAAGCAGTTGAAGATTAACTTGGTGA 324  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
DB 325 TTTTGAACCTGAGACAGAAAGAACCCAGAACAAATTTGCAAGAAACCAACATTAAG 384  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140  
DB 385 AAGAAAGAAATTGAGAAACTGCGAAGAAAGTGGCGCTGCATGAGCACTGGCTGCC 444  
QY 141 MetAsp 142  
DB 445 ATGGAT 450

RESULT 36  
LOCUS BU600854 828 bp mRNA linear EST 20-SEP-2002  
DEFINITION AGENCOURT\_10058678 NIH\_MGC\_142 Homo sapiens cDNA clone  
IMAGE:6454689 5', mRNA sequence.  
ACCESSION BU600854  
VERSION BU600854  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 828)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs@email.nih.gov](mailto:cgabs@email.nih.gov)  
Tissue Procurement: NCI

cDNA Library Preparation: Michael Brownstein Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: ULCM2671 row: f column: 02  
High quality sequence stop: 574.

## FEATURES

## source

1. 828  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="TM6S:6494689"  
/tissue\_type="mixed (pool of 40 RNAs)"  
/lab\_host="DH10B (T1-phage-resistant)"  
/clone\_lib="NIH MGC 142"  
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattacggc);  
Site 2: SfiI (ggccgcctggcc); Double-stranded cDNA was  
prepared from a pool of 40 cell line polyA+ RNAs (bladder  
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -  
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,  
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,  
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary  
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were  
used in cloning as follows:  
5'-AACGACGTGATCAAGCGACAGTGGCCATTACGCCGG-3' and  
5'-ATTCTAGAGCCGAGCGCGCCGACATG-AT(30)NN-3'. Full-length  
enriched library was constructed using the Clontech  
Creator SMART kit and size-selected to contain the >0.5 kb  
size fraction (other fractions present in NIH MGC 141).  
Library created in the laboratory of M. Brownstein (NIH,  
NIH). Note: this is a NIH\_MGC library."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.19e-148 Length: 828  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x BU60854 (1-828)

QY 1 MetGtAlAPrOthreuprOprOlaATpGInPrOphelLeuVAspHIsArgIleSer 20  
DB 68 ATGGGTCCTGACGCTGGCCCTGCTGGCGAGCCCTTTCTCAAGAGCACCGCATCTCT 127  
QY 21 ThrPhelYAsnTrpPrOphelengluGlyCySaIaCySthPrOgluArgMetAlaGlu 40  
DB 128 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTCGCCTGACCCCGAGCGGATGGCCGAG 187  
QY 41 AlaGlyPhelIleHisCySPrOthrgluengluPrOAspLeuAlaGlnCySPhPheYs 60  
DB 188 GCTGGCTTATCATCAGCCCACTAGAAAGGAGCAAGCTTGGCCCAAGTCTTCTTGGC 247  
QY 61 PhelYsGluLeuGluGlyTrpGluPrOAspAspAPPPrOleGluGluNHIsLysLysHis 80  
DB 248 TTCAAGAGACTGGAAGGCTGGAGGACAGATGACGACCCCATAGAGACATTAAGAT 307  
QY 81 SerSerGlyCySaIaPhelLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 308 TCGTCGGTTGGCGCTTCTTCTTCTCAAGAGAGGTTTGAAGATTAACTTGTGTGA 367

QY 101 PhelLeuYsLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
DB 368 TTTTGAACCTGACAGAGAGAAAGAGCCAAAGCAAAATTGCAAGAGAAACCAATTAAG 427  
QY 121 LysLysGluPheGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 140  
DB 428 AAGAAAGAAATTGAGGAACCTGCGAAGAAAGTGGCGGTCATCGACGAGCTGGCTGCC 487

QY 141 MetAsp 142  
DB 488 ATGAT 493

## RESULT 37

## LOCUS

AU126048 836 bp mRNA linear EST 01-AUG-2002  
AU126048 NT2RM4 Homo sapiens cDNA clone NT2RM4002626 5', mRNA  
sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## LOCUS

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 Db 175 GCTGGCTTCAATCCACTCCACCTGAGACGAGCACAATGGCCCGCTTCTTCCTGC 234  
 QY 61 PheLysGluLeuGluGlyTPGluPProAspAspAspProIleGluGluHisLysLysHis 80  
 Db 235 TTCAGAGAGCTGGAGAGCTGGAGCCAGATGACGACCCCATAGAGCACTATAAAGCAT 234  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 295 TCGTCCGGTTCGCTTCTTCTGTCAGAGCAGTTTGAAGATTACCTTGGTGA 354  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 Db 355 TTTTGAAGACTGACAGAGAAAGCCAGAACAAATTCGAAAGGAAACCAATAG 414  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140  
 Db 415 AAGAAAGATTGAGGAAAGCTGCCAGAAAGAGTGGCCGTCATCGAGCAGCTGGTGC 474  
 QY 141 MetAsp 142  
 Db 475 ATGGAT 480

RESULT 38  
 BQ438198 861 bp mRNA linear EST 24-MAY-2002  
 LOCUS BQ438198  
 DEFINITION AGENCOURT 7889851 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6175596  
 5', mRNA sequence.  
 ACCESSION BQ438198  
 VERSION BQ438198.1 GI:21177274  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 861)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgsabbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
 http://image.llnl.gov  
 Plate: LHM3550 row: j column: 13  
 High quality sequence stop: 573.  
 Location/Qualifiers  
 1..861  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6175596"  
 /tissue\_type="leiomyosarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 71"  
 /note="Organ: uterus; Vector: PCMV-SPORE; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.  
 Average insert size 2.1 Kb."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,23e-148 Length: 861  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x BQ438198 (1-861)  
 QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 5 ATGGGATGCCCCCAGCTTCCCTCCCTGCGAGCCCTTCTTCAAGACCAACCGCATCTCT 64  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 65 AATTTCAGAACTGGCCCTCTTCTTGAAGGCTCGCTGCACCCCGAGAGGATGCGCAG 124  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
 Db 125 GCTGGCTTCAATCCACTCCACCTGAGACGAGCACAATGGCCCGCTTCTTCCTGC 184  
 QY 61 PheLysGluLeuGluGlyTPGluPProAspAspAspProIleGluGluHisLysLysHis 80  
 Db 185 TTCAGAGAGCTGGAGAGCTGGAGCCAGATGACGACCCCATAGAGCACTATAAAGCAT 244  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 245 TCGTCCGGTTCGCTTCTTCTGTCAGAGCAGTTTGAAGATTACCTTGGTGA 304  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 Db 305 TTTTGAAGACTGACAGAGAAAGCCAGAACAAATTCGAAAGGAAACCAATAG 364  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140  
 Db 365 AAGAAAGATTGAGGAAAGCTGCCAGAAAGAGTGGCCGTCATCGAGCAGCTGGTGC 424  
 QY 141 MetAsp 142  
 Db 425 ATGGAT 430

RESULT 39  
 BQ882604 868 bp mRNA linear EST 16-AUG-2002  
 LOCUS BQ882604  
 DEFINITION AGENCOURT 6726226 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:6341045  
 5', mRNA sequence.  
 ACCESSION BQ882604  
 VERSION BQ882604.1 GI:22274612  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 868)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgsabbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
 http://image.llnl.gov  
 Plate: LHCW538 row: h column: 06  
 High quality sequence stop: 649.  
 Location/Qualifiers  
 1..868  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6341045"  
 /tissue\_type="neuroblastoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 47"  
 /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGACGAG (G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.:	1,248-148	Length:	868
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-09-690-825-34 (1-142) x B0882604 (1-868)

```
QY 1 MetGlyAlaProThrLeuProAlaTPGInProPheLeuLysAspHisArgIleSer 20
DB 44 ATGGGTCCTCATCAGCTGCGCCCTGCTGCGAGCCCTTTCTCAAGACACCGCATCTCT 103
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetLacIu 40
DB 104 ACATTCAAGAACTGGCGCTTCTTGAGAGGCTGCGCTGACCCCGAGCGGATGGCCGAG 163
QY 41 AlaGlyPheIleHisCysProThrGluLeuGluProAspLeuAlaGlnCysPhePheCys 60
DB 164 GCTGGCTTCATCAGCTGCGCCCTGCTGCGAGCCGAGCCGAGCTTGGCCAGTCTTCTTGC 223
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80
DB 224 TCAAGAGACTGGAAGGCTGGAGCCAGATGACGACCCCATPAGAGAACATPAAAAAGCAT 283
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB 284 TCGTCCGCTTGCCTTCTTCTTCTGTCAGAGCAGTTTGAAGATTAACTTGGTGA 343
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLysHis 120
DB 344 TTTTGAACCTGACGAGAAAGAGCCAAACAAATTGCCAAGAGAACCAACATTAAG 403
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaIa 140
DB 404 AAGAAAGATTGAGAGAACTGCGAAGAAAGTCCCGCTGCCATGACAGCTGGCTGCC 463
QY 141 MetAsp 142
DB 464 ATGAT 469
```

RESULT 40 BE883968 874 bp mRNA linear EST 20-OCT-2000  
LOCUS 601505487F1 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:3907388 5',  
DEFINITION mRNA sequence.

ACCESSION BE883968.1 GI:10332744  
VERSION BE883968.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 874) Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>  
Plate: LIA9717 row: n column: 21  
High quality sequence stop: 648.  
Location/Qualifiers  
1. 874  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3907388"  
/tissue\_type="leptomysarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb."

## FEATURES

## source

## ORIGIN

## Alignment Scores:

Pred. No.:	1,258-148	Length:	874
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-690-825-34 (1-142) x BE883968 (1-874)

```
QY 1 MetGlyAlaProThrLeuProAlaTPGInProPheLeuLysAspHisArgIleSer 20
DB 40 ATGGGTCCTCATCAGCTGCGCCCTGCTGCGAGCCCTTTCTCAAGACACCGCATCTCT 99
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetLacIu 40
DB 100 ACATTCAAGAACTGGCGCTTCTTGAGAGGCTGCGCTGACCCCGAGCGGATGGCCGAG 159
QY 41 AlaGlyPheIleHisCysProThrGluLeuGluProAspLeuAlaGlnCysPhePheCys 60
DB 160 GCTGGCTTCATCAGCTGCGCCCTGCTGCGAGCCGAGCATTTGGCCAGTCTTCTTCTGC 219
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80
DB 220 TCAAGAGACTGGAAGGCTGGAGCCAGATGACGACCCCATPAGAGAACATPAAAAAGCAT 279
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB 280 TCGTCCGCTTGCCTTCTTCTTCTGTCAGAGCAGTTTGAAGATTAACTTGGTGA 339
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLysHis 120
DB 340 TTTTGAACCTGACGAGAAAGAGCCAAACAAATTGCCAAGAGAACCAACATTAAG 399
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaIa 140
DB 400 AAGAAAGATTGAGAGAACTGCGAAGAAAGTCCCGCTGCCATGACAGCTGGCTGCC 459
QY 141 MetAsp 142
DB 460 ATGAT 465
```

RESULT 41 BE898710 875 bp mRNA linear EST 29-SEP-2000  
LOCUS 601681673F1 NIH\_MGC\_9 Homo sapiens cDNA clone IMAGE:3951667 5',  
DEFINITION mRNA sequence.

ACCESSION BE898710  
VERSION BE898710.1 GI:10365480  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 875) Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Incyte Genomics, Inc.



REFERENCE 1 (bases 1 to 875)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: [image.llnl.gov](http://image.llnl.gov)  
Plate: LNCM820 row: b column: 20  
High quality sequence stop: 709.  
Location/Qualifiers

FEATURES  
source

1..875  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3951667"  
/issue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1lb="NIH\_MGC\_9"  
/note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.25e-148 Length: 875  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-09-690-825-34 (1-142) x BE896710 (1-875)

QY 1 MetGtAlaProThrLeuProProAlaTrpGlnProPheLeuYsAspHisArgIleSer 20  
DB 6 ATGGGTGGCCGAGCGTTCCCTCCCTGCGAGCGCTTCTCAAGACACCGCATCTCT 65  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 66 ACATTCAGAACTGGCCCTTCTTGGAGGGCTGGCGCTCCACCCCGAGCGGATGAGCGAG 125  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 126 GCTGCTTCATCCATGCGCCCTCCCTGAGACGAGCGAGCTTGGCCCATGTTCTTCGCG 185  
QY 61 PheLysGluLeuGluGlyTTPGluProAspAspAspProIleGluGluHisLysLysHis 80  
DB 186 TTCAGAGGAGCTGGAGGCTGGAGGCGAGTGAACGCCCAAGAGAAACATTAAGAGAT 245  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 246 TGGTCCGGTGGCGCTTCTTCTCTGTCAGAAAGCGAGTTGAAGATTAACTTCGTGAA 305  
QY 101 PheLeuYsLeuAspArgGluArgAlaLysAsnLysIleAlaLysGlnThrAsnLys 120  
DB 306 TTTTCAAACTGAGACGAGAAAGAGCAAAATTTGCAGAAAGAAACCAACATATAG 365  
QY 121 LysLysGluPheGluGluThrAlaLysLysValaArgAlaIleGluGluLeuAlaIa 140  
DB 366 AAGAAAGAAATTTGAGAAACTGAGAAAGAGGCGCGCTGCGATCGAGAGCTGCTGCC 425  
QY 141 MetAsp 142  
|||||

DB 426 ATGCAT 431

RESULT 42  
LOCUS BG528273 875 bp mRNA linear EST 03-APR-2001  
DEFINITION 602557679P1 NIH\_MGC\_59 Homo sapiens CDNA clone IMAGE:4686547 5',  
mRNA sequence.  
ACCESSION BG528273  
VERSION BG528273.1 GI:13519810  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 875)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC

CDNA Library Preparation: CLONETECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Plate: LNCM1497 row: n column: 20  
High quality sequence stop: 692.  
Location/Qualifiers

FEATURES  
source

1..875  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4686547"  
/issue\_type="mucoepidermoid carcinoma"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_1lb="NIH\_MGC\_59"  
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccgctcggc); Site 2: SfiI (ggccatcggc); SfiI (ggccgctcggc); Double-stranded CDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGAGGCGGAGCGCGGAGCATG-dT(30)BN-3' (where B = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.25e-148 Length: 875  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-09-690-825-34 (1-142) x BG528273 (1-875)

QY 1 MetGtAlaProThrLeuProProAlaTrpGlnProPheLeuYsAspHisArgIleSer 20  
DB 46 ATGGGTGGCCGAGCGTTCCCTCCCTGCGAGCGCTTCTCAAGACACCGCATCTCT 105  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 106 ACATTCAGAACTGGCCCTTCTTGGAGGGCTGGCGCTGCAACCCGAGGAGGAGGCGGAG 165  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 166 GCTGCTTCATCCATGCGCCCTCCCTGAGACGAGCGAGCTTGGCCCATGTTCTTCTGCG 225



## ORIGIN

of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

## Alignment Scores:

Pred. No.:	1.26e-148	Length:	883
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-09-690-825-34 (1-142) x BQ650551 (1-883)

QY 1 MetGlyAlaProThrLeuProProAlaTTPGlnProPheLeuLysAspHisArgIleSer 20

Db 29 ATGGGTGCCCCGAGCTTGCCCTGCTGGCAGCCCTTTCTCAAGACACACCGCATCTCT 88

QY 21 ThrPheLysAsnTTPProPheLeuGluGlyCysValaCysThrProGluArgMetAlaGlu 40

Db 89 ACATTCAAGAACTGCGCCCTTCTTGAGAGGCTGCGCCCTGCACCCCGAGCGGATGCGGAG 148

QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60

Db 149 GCTGGCTTCATCCACTGCCCCACTGAGACAGACACTTGCGCCCACTGTTCTTCTGTC 208

QY 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluLysIleLysLysHis 80

Db 209 TTCAAGAGAGCTGGAAGCTGGAGACAGATGACGCCCATAGAGAACTAAAGACAT 268

QY 81 SerSerGlyCysAlaPheLeuSerValIleLysGlnPheGluGluLeuThrLeuGlyGlu 100

Db 269 TCGTCCGCTGCGCTTCTTCTTCTGTCAGAGAGCGCTTTGAAGAAATTACCCCTTGAGAA 328

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120

Db 329 TTTTGAAGAACTGACAGAGAAAGAGCCAAAGAAATTTGCAAGAAACCAACATTAAG 388

QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140

Db 389 AAGAAAGAAATTTGAGAAACTGCGAAGAAAGTGGCGCTGCATCGAGCACTGGCTGCC 448

QY 141 MetAsp 142

Db 449 ATGGAT 454

RESULT 45 BQ227378 888 bp mRNA linear EST 02-MAY-2002

LOCUS AGENCOURT 7592800 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6050370

DEFINITION 5', mRNA sequence.

ACCESSION BQ227378

VERSION BQ227378.1 GI:20408778

KEYWORDS EST.

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 888)

AUTHORS NIH-MGC http://mgs.nci.nih.gov/.

TITLE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: ATCC/DC/DTP  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov

Plate: L1AM13302 row: h column: 19  
High quality sequence stop: 554.  
Location/Qualifiers

## FEATURES

## source

1..888  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6050370"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_72"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site\_1: NotI,  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

## ORIGIN

## Alignment Scores:

Pred. No.:	1.26e-148	Length:	888
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-09-690-825-34 (1-142) x BQ227378 (1-888)

QY 1 MetGlyAlaProThrLeuProProAlaTTPGlnProPheLeuLysAspHisArgIleSer 20

Db 46 ATGGGTGCCCCGAGCTTGCCCTGCTGGCAGCCCTTTCTCAAGACACACCGCATCTCT 105

QY 21 ThrPheLysAsnTTPProPheLeuGluGlyCysValaCysThrProGluArgMetAlaGlu 40

Db 106 ACATTCAAGAACTGCGCCCTTCTTGAGAGGCTGCGCCCTGCACCCCGAGCGGATGCGGAG 165

QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60

Db 166 GCTGGCTTCATCCACTGCCCCACTGAGACAGACACTTGAGCCCACTGTTCTTCTGTC 225

QY 61 PheLysGluLeuGluGlyTTPGluProAspAspProIleGluLysIleLysLysHis 80

Db 226 TTCAAGAGAGCTGGAAGCTGGAGACAGATGACGCCCATAGAGAACTAAAGACAT 285

QY 81 SerSerGlyCysAlaPheLeuSerValIleLysGlnPheGluGluLeuThrLeuGlyGlu 100

Db 286 TCGTCCGCTGCGCTTCTTCTTCTGTCAGAGAGCGCTTTGAAGAAATTACCCCTTGAGAA 345

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120

Db 346 TTTTGAAGAACTGACAGAGAAAGAGCCAAAGAAATTTGCAAGAAACCAACATTAAG 405

QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140

Db 406 AAGAAAGAAATTTGAGAAACTGCGAAGAAAGTGGCGCTGCATCGAGCACTGGCTGCC 465

QY 141 MetAsp 142

Db 466 ATGGAT 471

RESULT 46 BQ250436 889 bp mRNA linear EST 13-FEB-2001

LOCUS 602362594F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:4471055 5',

DEFINITION mRNA sequence.

ACCESSION BQ250436

VERSION BQ250436.1 GI:12760252

KEYWORDS EST.

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 889)

AUTHORS NIH-MGC http://mgs.nci.nih.gov/.

TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

JOURNAL

COMMENT

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILNI at:  
http://image.jnl.gov  
Plate: LLM10289 row: k column: 24  
High quality sequence stop: 682.

## FEATURES

source

Location/Qualifiers  
1..889  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4471055"  
/tissue\_type="adenoecarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NH\_MGC\_90"  
/note="Organ: liver; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.27e-148 Length: 889  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-690-825-34 (1-142) x BG250436 (1-889)

QY 1 MetGtAlAPrOThrLeuPrOProAlATrPgiNPrOphLeuLysAspHisArgIleSer 20  
DB 15 ATGGGTCGCCCAAGCTGGCCCTGGCTGGCAGCCCTTCTCAAGAGCACCGCATCTCT 74  
QY 21 ThrPhelYAsnTrpProPhelengLugLYCySaIaCyStrPrOgiUArgMetAlaGlu 40  
DB 75 ACATTCAAGAACTGGGCGCTTCTTGAGAGGCTCGCGCTGACCCCGAGCGATGGCCGAG 134  
QY 41 AlaGlyPheIleHisCysEsrProThrGluAengLuprOAspLeuAlaGlnCysPhePheCys 60  
DB 135 GCTGGCTTCAATCCACTGCCCCCACTAGAAAGCAAGCAAGCTGGCCAGTGTCTTCTGC 194  
QY 61 PheLysGluLeuGluGlyTrpGluPrOAspAspProIleGluGlnHisLysLysHis 80  
DB 195 TTCAGAGAGCTGGAAGGCTGGAGCAATGACACCCCAATGAGGAACATTAAGAAGCAT 254  
QY 81 SerSerGlyCyAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 255 TCGTCCGCTTGGCTTCTTCTTCTGTCAGAAAGCGATTGAAGATTAACTTGGTGA 314  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
DB 315 TTTTGAAGCTGACAGAGAAAGACCAAGAAATTCGAAGGAACCAACATTAAG 374  
QY 121 LysLysGluPheGluGlnThrAlaLysLysValaArgArgAlaIleGluGlnLeuAla 140  
DB 375 AAGAAGAATTGAGGAAGAACTGCAAGAAAGTGCCTGCATCGACACACTGCTGCC 434  
QY 141 MetAsp 142  
DB 435 ATGGAT 440

RESULT 47  
BX425472

LOCUS BX425472 896 bp mRNA linear EST 15-MAY-2003  
DEFINITION BX425472 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone  
ACCESSION CL08B0152E06 5-PRIME, mRNA sequence.  
VERSION BX425472  
KEYWORDS BX425472.1 GI:30778469  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 896)  
AUTHORS Li W.B., Gruber, C., Jesse, J. and Polyes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 8555.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CL08B0152E06R1acluster=8555.r. Contact :  
Feng Liang Email: fliang@life.techn.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CL08B0152E06R1.

## FEATURES

source

Location/Qualifiers  
1..896  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CL08B0152E06"  
/tissue\_type="NEUROBLASTOMA"  
/clone\_id="Homo sapiens NEUROBLASTOMA"  
/note="Vector: pCMVSPORT6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.27e-148 Length: 896  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x BX425472 (1-896)

QY 1 MetGtAlAPrOThrLeuPrOProAlATrPgiNPrOphLeuLysAspHisArgIleSer 20  
DB 75 ATGGGTCGCCCAAGCTGGCCCTGGCTGGCAGCCCTTCTCAAGAGCACCGCATCTCT 134  
QY 21 ThrPhelYAsnTrpProPhelengLugLYCySaIaCyStrPrOgiUArgMetAlaGlu 40  
DB 135 ACATTCAAGAACTGGGCGCTTCTTGAGAGGCTCGCGCTGACCCCGAGCGATGGCCGAG 194  
QY 41 AlaGlyPheIleHisCysEsrProThrGluAengLuprOAspLeuAlaGlnCysPhePheCys 60  
DB 195 GCTGGCTTCAATCCACTGCCCCCACTAGAAAGCAAGCAAGCTGGCCAGTGTCTTCTGC 254  
QY 61 PheLysGluLeuGluGlyTrpGluPrOAspAspProIleGluGlnHisLysLysHis 80  
DB 255 TTCAGAGAGCTGGAAGGCTGGAGCAATGACACCCCAATGAGGAACATTAAGAAGCAT 314  
QY 81 SerSerGlyCyAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 315 TCGTCCGCTTGGCTTCTTCTTCTGTCAGAAAGCGATTGAAGATTAACTTGGTGA 374  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120

Db 375 TTTTGAACCTGCACAGAAAGCCAGAACCAATTCGAAAGAACCAATTAAG 434

Qy 121 LysylsGluPheGluGluThrAlaLysValArgArgAlaIleGluGluLeuAla 140

Db 435 AAGAAAGAAATTGAGAACTGCAGAAAGAAAGTGCCTGCATCGAGCAGTGGTGC 494

Qy 141 MetAsp 142

Db 495 ATGGAT 500

RESULT 48

BF972840 908 bp mRNA linear EST 22-JAN-2001

LOCUS 602241117F1 NIH\_MGC\_46 Homo sapiens cDNA clone IMAGE:4329509 5',

DEFINITION mRNA sequence.

ACCESSION BF972840

VERSION BF972840.1 GI:12340055

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS 1 (bases 1 to 908)

TITLE NIH-MGC http://mgs.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapds-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov

Plate: L10M191 row: b column: 06

High quality sequence stop: 687.

Location/Qualifiers

1..908

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4329509"

/tissue\_type="leiomyosarcoma cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 46"

/note="Organ: uterus; Vector: POTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 1.29e-148 Length: 908

Score: 142.00 Matches: 142

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 10 Gaps: 0

US-09-690-825-34 (1-142) x BF972840 (1-908)

Qy 1 MetGlyAlaProThrLeuProProAlaITpGlnProPheLeuLysAspHisArgIleSer 20

Db 13 ATGGAGGCCCGCGAGCTTGGCCCTGCGACGACCTTTCTCAAGACCAACGCACTCTT 72

Qy 21 ThrPheLysAsnTrpProPheLeuGluGlyValaCysThrProGluArgValaGlu 40

Db 73 ACATTCAGAACTGCCTCTTCTTGGAGGCTGCGCCTGCACCCCGAGCGGATGGCCGAG 132

Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGluCysPhePheCys 60

Db 133 GGTGGCTTCATCTACCTGCCCCCTAGAACGAGCCACTTGCCCATGTTCTTCTTCC 192

Qy 61 PheLysGluLeuGluGlyTyrGluProAspAspAspProIleGluHisLysLysHis 80

Db 193 TTCAGAGAGCTGGAAGGCTGGAGCCAGATGACACCCATAGAGAAATMAAAGCAT 252

Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100

Db 253 TCGTCGGTGGCTTCTTCTTCTGCAAGAGAGGTTGAAGAAATTAACCTTGATGA 312

Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120

Db 313 TTTTGAACCTGCACAGAAAGAACCCAAAGAAATTCGAAAGAAACCAATTAAG 372

Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAla 140

Db 373 AAGAAAGAAATTGAGAAACTGCAGAAAGAAAGTGCCTGCATGAGCAGCTGGCTGC 432

Qy 141 MetAsp 142

Db 433 ATGGAT 438

RESULT 49

BG110508 915 bp mRNA linear EST 30-JAN-2001

LOCUS 602278934F1 NIH\_MGC\_86 Homo sapiens cDNA clone IMAGE:4366425 5',

DEFINITION mRNA sequence.

ACCESSION BG110508

VERSION BG110508.1 GI:12604014

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS 1 (bases 1 to 915)

TITLE NIH-MGC http://mgs.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapds-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov

Plate: L10M10017 row: d column: 10

High quality sequence stop: 606.

Location/Qualifiers

1..915

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4366425"

/tissue\_type="osteosarcoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 86"

/note="Organ: bone; Vector: PCMV-Spore6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.53 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 1.3e-148 Length: 915

Score: 142.00 Matches: 142

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-690-825-34 (1-142) x BG110508 (1-915)

QY 1 MetG1YAlaProThreupProAlaTPGInProPhelenuysAsph1sarq1leser 20  
DB 31 ATGGGTCCTCAGTGGCCCTGCTGCAAGACCTTCTCAAGACCAACGCACTCT 90  
QY 21 ThrPhelysAsnTrpProPhelenuyG1YCyAlaCysThrProGluArgMetAlaGlu 40  
DB 91 ACATTGAGAACTGGCCCTTCTTGAGGGCTGGCTGCACTGCACTGCACTGCACTG 150  
QY 41 AlaGlyPhe1LeHisCysProThrg1uaenGluProAsp1leuAlaGlnCysPhePheCys 60  
DB 151 GCTGGCTTCATCAGTGGCCCTGCTGCAAGACCACTGCACTGCACTGCACTGCACT 210  
QY 61 Phe1ySGluLeuGlu1YTrpGluProAspAspPro1leGluGluHis1yS1ySHis 80  
DB 211 TTCAAGAGCTGAGAGCTGGAGCCAGATGACGACCCCATAGAGAAATMAAAGAT 270  
QY 81 SerSerG1YCyAlaPhe1euserVal1yS1ySGlnPheG1u1leuThre1uG1u 100  
DB 271 TCGTCCGGTGGCTTCTTCTTCTGCAAGACATTTGAAGATTAACCTTGTGTA 330  
QY 101 Phe1eu1ySGluAspArgG1uArgAla1ySAsn1yS1leAla1ySGluThrasn1yS 120  
DB 331 TTTTAAACTGACAGAGAAAGCCAAAGCAAAATTGCAAGAGAAACCAACAATAG 390  
QY 121 Lys1ySGluPheGlu1uThra1a1yS1ySVal1ArgArgAla1leG1u1leuAla1a 140  
DB 391 AAGAAAGAAATTGAGAAACTGCGAAGAAAGTCCCGCTGCATGACAGCTGCTGCC 450  
QY 141 MetAsp 142  
DB 451 ATGGAT 456

RESULT 50  
BI760504 924 bp mRNA linear EST 25-SEP-2001  
LOCUS 603045042F1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5185114 5',  
DEFINITION mRNA sequence.  
ACCESSION BI760504  
VERSION BI760504.1 GI:15752082  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LML at:  
http://image.llnl.gov  
Plate: LML11462 row: d column: 11  
High quality sequence stop: 862.  
Location/Qualifiers  
1..924  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5185114"  
/lab\_host="DH10B"  
/clone\_11b="NIH\_MGC\_116"

FEATURES  
source

/note="Organ: pooled colon, kidney, stomach; Vector: PCMV-SPORT6; Site 1: Nct1; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 clones, age 26 yo male, 49 yo female, 71 yo male colon, 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 1,31e-148 Length: 924  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-690-825-34 (1-142) x BI760504 (1-924)

QY 1 MetG1YAlaProThreupProAlaTPGInProPhelenuysAsph1sarq1leser 20  
DB 10 ATGGGTCCTCAGTGGCCCTGCTGCAAGACCTTCTCAAGACCAACGCACTCT 69  
QY 21 ThrPhelysAsnTrpProPhelenuyG1YCyAlaCysThrProGluArgMetAlaGlu 40  
DB 70 ACATTGAGAACTGGCCCTTCTTGAGGGCTGGCTGCACTGCACTGCACTGCACTG 129  
QY 41 AlaGlyPhe1LeHisCysProThrg1uaenGluProAsp1leuAlaGlnCysPhePheCys 60  
DB 130 GCTGGCTTCATCAGTGGCCCTGCTGCAAGACCACTGCACTGCACTGCACTGCACT 189  
QY 61 Phe1ySGluLeuGlu1YTrpGluProAspAspPro1leGluGluHis1yS1ySHis 80  
DB 190 TTCAAGAGCTGAGAGCTGGAGCCAGATGACGACCCCATAGAGAAATMAAAGAT 249  
QY 81 SerSerG1YCyAlaPhe1euserVal1yS1ySGlnPheG1u1leuThre1uG1u 100  
DB 250 TCGTCCGGTGGCTTCTTCTTCTGCAAGACATTTGAAGATTAACCTTGTGTA 309  
QY 101 Phe1eu1ySGluAspArgG1uArgAla1ySAsn1yS1leAla1ySGluThrasn1yS 120  
DB 310 TTTTGAACCTGACAGAGAAAGCCAAAGCAAAATTGCAAGAGAAACCAACAATAG 369  
QY 121 Lys1ySGluPheGlu1uThra1a1yS1ySVal1ArgArgAla1leG1u1leuAla1a 140  
DB 370 AAGAAAGAAATTGAGAGAAACTGCGAAGAAAGTCCCGCTGCATGACAGCTGCTGCC 429  
QY 141 MetAsp 142  
DB 430 ATGGAT 435

RESULT 51  
BQ959492 933 bp mRNA linear EST 21-AUG-2002  
LOCUS BQ959492  
DEFINITION AGENCOURT 8922878 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6473127  
5' RNA Sequence.  
ACCESSION BQ959492  
VERSION BQ959492.1 GI:22374970  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LLM14008 row: C column: 16  
High quality sequence stop: 630.  
Location/Qualifiers

## FEATURES

source

1..933  
/organism="Homo sapiens"  
/mol\_type="rRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6473127"  
/tissue\_type="leiomysarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 71"  
/note="Organ: uterus; Vector: PCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.  
Average insert size 2.1 kb."

## ORIGIN

## Alignment Scores:

Score:	1.32e-148	Length:	933
Percent Similarity:	142.00	Matches:	142
Best Local Similarity:	100.00%	Conservative:	0
Query Match:	100.00%	Mismatches:	0
		Indels:	0
		Gaps:	0

US-09-690-825-34 (1-142) x BQ959492 (1-933)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 25 ATGGGTGCCCCGACGCTTCCCTCCCTGCGCCCTTCTCAAGACACCGACCTCT 84  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 85 ACATTCAAGAACTGGCCCTTCTTGAGGGCTGCGCTGCACCCCGAGCGATGGCCGAG 144  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 145 GCTGCTTCATCCACTGCCCCCTAGAACGACGACCTGGCCCACTGTTCTTCTGCG 204  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisLysLysHis 80  
DB 205 TTCAAGAGCTGGAAGCGCTGGAGCAGATGACGCCCATAGAGAACTAAAGAT 264  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 265 TCGTCCGGTGGCTTCTTCTTCTCAAGACGATTGAAGATTAAACCTTGTGAA 324  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
DB 325 TTTTGAACCTGACAGAGAAAGACGACAAACAAATTGCAAGAAACCAACAATAG 384  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAla 140  
DB 385 AAGAAAGATTGAGAGAACTGCGAAGAAAGTGGCCCGCATGACAGCTGGCTGCC 444  
QY 141 MetAsp 142  
DB 445 ATGAT 450

RESULT 52  
BF982118 934 bp mRNA linear EST 23-JAN-2001  
LOCUS BF982118  
DEFINITION 602308926f1 NIH\_MGC\_88 Homo sapiens CDNA clone IMAGE:4400190 5',  
mRNA sequence.  
ACCESSION BF982118  
VERSION BF982118.1 GI:12384930  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 934)  
TITLE NIH-MGC <http://mc.ncbi.nih.gov/>.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
published (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LLM10105 row: C column: 07  
High quality sequence stop: 630.  
Location/Qualifiers

## FEATURES

source

1..934  
/organism="Homo sapiens"  
/mol\_type="rRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4400190"  
/tissue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 88"  
/note="Organ: small\_intestine; Vector: PCMV-SPORT6;  
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;  
oligo-dt primed. Average insert size 1.767 kb. Library  
enriched for full-length clones and constructed by Life  
Technologies. Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Score:	1.32e-148	Length:	934
Percent Similarity:	142.00	Matches:	142
Best Local Similarity:	100.00%	Conservative:	0
Query Match:	100.00%	Mismatches:	0
		Indels:	0
		Gaps:	0

US-09-690-825-34 (1-142) x BF982118 (1-934)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 13 ATGGGTGCCCCGACGCTTCCCTCCCTGCGCCCTTCTCAAGACACCGACCTCT 72  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 73 ACATTCAAGAACTGGCCCTTCTTGAGGGCTGCGCTGCACCCCGAGCGATGGCCGAG 132  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 133 GCTGCTTCATCCACTGCCCCCTAGAACGACGACCACTGGCCCACTGTTCTTCTGCG 192  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisLysLysHis 80  
DB 193 TTCAAGAGCTGGAAGCGCTGGAGCAGATGACGCCCATAGAGAACTAAAGAT 252  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 253 TCGTCCGGTGGCTTCTTCTTCTCAAGACGATTGAAGATTAAACCTTGTGAA 312  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
DB 313 TTTTGAACCTGACAGAGAAAGACGACAAACAAATTGCAAGAAACCAACAATAG 372  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAla 140  
DB 373 AAGAAAGATTGAGAGAACTGCGAAGAAAGTGGCCCGCATGACAGCTGGCTGCC 432

QY 141 Met-Asp 142  
Db 433 ATGGAT 438  
RESULT 53  
BE883927  
LOCUS BE883927 935 bp mRNA linear EST 20-OCT-2000  
DEFINITION 601505442F1 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:3907138 5',  
mRNA sequence.  
ACCESSION BE883927  
VERSION BE883927  
KEYWORDS EST.  
SOURCE BE883927.1 GI:10332703  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 935)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM9717 row: c column: 11  
Location/Qualifiers  
1. 935  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3907138"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2.1 kb."  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,32e-148 Length: 935  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0  
US-09-690-825-34 (1-142) x BE883927 (1-935)  
ORIGIN  
QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 5 ATGGGGCCCCGACGCTGGCCCCCTGCGACGCTTTCACAGACACACGACATCTCT 64  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 65 AACTTCAGAACTGGCCCTCTCTTGGAGGCGCTCGCTCGACCCCGAGCGGATGGCCAG 124  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
Db 125 GCTGGCTTCATCCACTGCGCCCACTGAGAACGACGACGACTGGCCCACTGTTCTTCG 184  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspArgPheIleGluGluHisLysLysHis 80  
Db 185 TTCAGAGAACTGGAAGGCTGGAGGCAAGTGCAGCCCACTGAGAGAAACATPAAAAAGCAT 244  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100

Db 245 TCGTCGCGTTCGCGCTTTCCTTCTGTCACAGACGAGTTGAGATTACCTTGATGA 304  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluTrpAsnLys 120  
Db 305 TTTTGAACCTGCAGACAGAAAGCCACAGAACAAATTGCCAAGAACCAATTAAG 364  
QY 121 LysLysGluPheGluGluTrpAlaLysLysValArgArgAlaIleGluGlnLeuAla 140  
Db 365 AAGAAAGAAATTGAGAACTGCGAAGAAAGTCCGCTGCATCGACGAGCTGCTGCC 424  
QY 141 Met-Asp 142  
Db 425 ATGGAT 430  
RESULT 54  
BF981768 943 bp mRNA linear EST 23-JAN-2001  
LOCUS BF981768  
DEFINITION 602305433F1 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:4396689 5',  
mRNA sequence.  
ACCESSION BF981768  
VERSION BF981768  
KEYWORDS EST.  
SOURCE BF981768.1 GI:12384580  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 943)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10096 row: a column: 10  
Location/Qualifiers  
1. 943  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4396689"  
/tissue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 88"  
/note="Organ: small intestine; Vector: pCMV-SPORT6;  
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;  
Oligo-dT primed. Average insert size 1.767 kb. Library  
enriched for full-length clones and constructed by Life  
Technologies. Note: this is a NIH\_MGC Library."  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,33e-148 Length: 943  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0  
US-09-690-825-34 (1-142) x BF981768 (1-943)  
ORIGIN  
QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 9 ATGGGGCCCCGACGCTGGCCCCCTGCGACGCTTTCACAGACACACGACATCTCT 68  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40



	Dc	69	ACATTCAAGAACTGGCCCTTCTTGAGAAGGCGGCCCTGCACCCTGGAGCGCATGCGCAG	128
Qy		41	AlaGlyPheIleHisCySProThgAlaAsnGluProAspLeuAlaGlnCySPhePheCys	60
Dd		129	GCTGGCTTCATCCACTGCCCCCATGAGACGACGACGTGGCCAGTGTTCCTTCCTGC	188
Qy		61	PheYsgIuLeuEngJugLyTrpGluProAspAspAPerIolegluJuhIslyLysHis	80
Dd		189	TTCAGAGAGCTGGAAAGGTGGGACCATGACGCCCATAGAGGAACATPAAAAAAGCAT	248
Qy		81	SerSerGlyCysAlaPheLeuSerValIlySylsgInPhegluGluLeuThrLeuGlyGlu	100
Dd		249	TGCGTCGGGTGCGCTTCCTTCCTTCGTCTCAAGAGACAGTTGAAGATTAACTTGCGTGA	308
Qy		101	PheLeuYlsLeuAspArgAlaArgAlaIlysAsnYlsIleAlaIysGluThrAsnAspLys	120
Dd		309	TTTTTGAAACTGCAGACAGAAAGAACCCAGAACAAATGCGAAGAAAGAACCAACAATAAG	368
Dd		369	AAGAAGAAATTGAGAAAACGCGAAGAAAGTGGCGCTGCATCGACAGCACTGGCTGCC	428
Qy		141	MetAsp	142
Dd		429	ATGGAT	434
RESULT 55			/	
LOCUS	B0929774	943 bp	mRNA	linear EST 20-AUG-2002
DEFINITION	AGENCOURT_8950028 NIH_MGC_71 Homo sapiens CDNA clone IMAGE:6471370			
ACCESSION	B0929774			
VERSION	B0929774.1	GI:22344805		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiade; Homo.			
TITLE	NIH-MGC http://mgc.nci.nih.gov/.			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgsab@remail.nih.gov Issue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLNL4003 row: j column: 11 High quality sequence stop: 636. Location/Qualifiers 1..943 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6471370" /tissue_type="leiomyosarcoma" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 71" /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."			
FEATURES	source			
ORIGIN				
Alignment Scores:				
Pred. No.:	1..33e-148	Length:	943	
Score:	142.00	Matches:	142	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	

US-09-690-825-34 (1-142) x B0929774 (1-943)

DB: 13 gaps: 0

QY 1 MetGlyAlaProthrLeuProProlaTrpGlnProPheLeuYsaPshSargIleSer 20

Db 32 ATGGGATGCCCGAGCGTTTCCCCCTGCTGGCAGCCCTTTCTCAAGACACACCGCATCTCT 91

QY 21 ThrPheYsaAntpProPheLeuGluYcysAlaCysThrProGluArgMetAlaGlu 40

Db 92 ACATTCAAGAACTGGCCCTCTTGAGAGGCTGCCTCCGACCCCGGAGCGGATGCCGAG 151

QY 41 AlaGlyhellehscYseProThrGluAsnGlnProAspLeuIaGlnCysPhePheCys 60

Db 152 GCTGGCTTCATCCACTGCCCATCTGACAGACAGCACTGGCCCATGTTTCTTCTTGC 211

QY 61 PheYsgIuLeuGluGlyTrpGluProAspAspProIleGluIuHisIysIysHis 80

Db 212 TTCACAGAGCTCGAAGGCTGGAGCCAGAGACACCCCATAGAGGAACATTAAGAT 271

QY 81 SerSerGlyCysAlaPheLeuSerValIysGlnIheGluGluLeuGlyGlu 100

Db 272 TCGCCGATGCCGCTTCTTCTTCTGTCAGAGAGCACTTGAAAGATTAACTTGGTGA 331

QY 101 PheLeuYsaLeuAspArgGluArgAlaIysAsnIysIleAlaIysGluThrAsnAluIys 120

Db 332 TTTTGAACCTGACAGAGAAAGGCCAGAGACAAATTGCAAGAGAAACCAACATTAAG 391

QY 121 LysIysGluPheGluGluThrAlaIysIysValArgArgAlaIleGluGlnLeuAla 140

Db 392 AAGAAGATTGAGAAACTGCCAGAGAAATGCCGCGTGCATCGAGCACTGGCTGCC 451

QY 141 MetAsp 142

Db 452 ATGGAT 457

RESULT 56

B0649501

LOCUS

DEFINITION

AGENCOURT\_8492482 NIH\_MGC\_100 Homo sapiens cDNA clone IMAGE:6296575

5', mRNA sequence.

BO649501

BO649501.1 GI:21773673

EST.

Source

ORGANISM

Homo sapiens (human)

Homosapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 962)

NIH-MGC <http://mgs.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strusberg, Ph.D.

Email: [cgasbs-remail.nih.gov](mailto:cgasbs-remail.nih.gov)

Tissue Procurement: CGAP (Stanford)

cDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

CDNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: L10CM2504 row: k column: 08

High quality sequence stop: 618.

Location/Qualifiers

1..962

organism="Homo sapiens"

/mol\_type="RNA"

/db\_xref="taxon:9606"

/clone\_image="6296575"

/tissue\_type="hepatocellular carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 100"

/note="Organ: liver; Vector: pOTB7, Site\_1: XhoI; Site\_2:

FEATURES

source

ECORI: cDNA made by oligo-dT priming. Directionally cloned into EcorI/XhoI sites using the following 5' adaptor: GGCAACAG(G). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC library."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.35e-148  
Score: 142.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 13

Length: 962  
Matches: 142  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-690-825-34 (1-142) x BQ649501 (1-962)

QY 1 MetGlyAlaProThrLeuProProAlaTyrGlnProPheLeuLysAspHisArgIleSer 20  
Db 13 ATGGGCGCCCGACGCTGGCCCTGCTGGACGCCCTTTCTCAAGACCAACCGCATCTCT 72  
QY 21 ThrPheLysAsnTyrProPheLeuGluGlyCyAlaCysThrProGluArgMetAlaGlu 40  
Db 73 ACATTCAAGAACTGGCCCTTCTTGAGAGCGCTGCGCTGCACCCCGAGCGGATGCGCGAG 132  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlyCysPhePheCys 60  
Db 133 GCTGGCTTCATCCACCTGCCCTCACTGAAACGAGCCAGACTTGAGCCAGTGTCTTCTGCG 192  
QY 61 PheLysGluLeuGluGlyTyrGluProAspAspProIleGluGluHisLysLysHis 80  
Db 193 TTCAGAGAGCTGGAAGCTGGAGCGAGATGACGACCCCATAGAGAAATTAAGCAT 252  
QY 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluGluLeuThrLeuGlyGlu 100  
Db 253 TCGTCCGGTTCGCTTCTTCTGTCAGAGGAGCTTGAAGAAATTAACCTTGCTGAA 312  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
Db 313 TTTTAAACTGGACAGAGAAAGAGCCAAATTCGAAAGAAACCAACATTAAG 372  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAla 140  
Db 373 AAGAAAGAAATTGAGGAAACTGCGAAGAAAGTGGCCGTCATGACGAGCTGGCTGCC 432  
QY 141 MetAsp 142  
Db 433 ATGGAT 438

RESULT 57  
B1114304 963 bp mRNA linear EST 26-JUN-2001  
LOCUS 602862321F1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:5021644 5',  
DEFINITION mRNA sequence.

ACCESSION B1114304  
VERSION B1114304  
KEYWORDS B1114304.1 GI:14565205  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 963)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: LNCM1840 row: 1 column: 05  
High quality sequence stop: 738.  
location/Qualifiers

## FEATURES

## SOURCE

1..963  
/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5021644"

/issue\_type="rhabdomyosarcoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_1lb="NIH\_MGC\_17"

/note="Organ: muscle; Vector: pOTB7, Site\_1: EcorI;  
Site\_2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcorI/XhoI sites using the  
following 5' adaptor: GGCAACAG(G). Size selected >500bp  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.35e-148  
Score: 142.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 12

Length: 963  
Matches: 142  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-690-825-34 (1-142) x B1114304 (1-963)

QY 1 MetGlyAlaProThrLeuProProAlaTyrGlnProPheLeuLysAspHisArgIleSer 20  
Db 34 ATGGGCGCCCGACGCTGGCCCTGCTGGACGCCCTTTCTCAAGACCAACCGCATCTCT 93  
QY 21 ThrPheLysAsnTyrProPheLeuGluGlyCyAlaCysThrProGluArgMetAlaGlu 40  
Db 94 ACATTCAAGAACTGGCCCTTCTTGAGAGCGCTGCGCTGCACCCCGAGCGGATGCGCGAG 153  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlyCysPhePheCys 60  
Db 154 GCTGGCTTCATCCACCTGCCCTCACTGAAACGAGCCAGACTTGAGCCAGTGTCTTCTGCC 213  
QY 61 PheLysGluLeuGluGlyTyrGluProAspAspProIleGluGluHisLysLysHis 80  
Db 214 TTCAGAGAGCTGGAAGCTGGAGCGAGATGACGACCCCATAGAGAAATTAAGCAT 273  
QY 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluGluLeuThrLeuGlyGlu 100  
Db 274 TCGTCCGGTTCGCTTCTTCTGTCAGAGGAGCTTGAAGAAATTAACCTTGCTGAA 333  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
Db 334 TTTTAAACTGGACAGAGAAAGAGCCAAATTCGAAAGAAACCAACATTAAG 393  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAla 140  
Db 394 AAGAAAGAAATTGAGGAAACTGCGAAGAAAGTGGCCGTCATGACGAGCTGGCTGCC 453  
QY 141 MetAsp 142  
Db 454 ATGGAT 459

RESULT 58  
B1114304 965 bp mRNA linear EST 12-SEP-2002  
LOCUS 602862321F1 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:6341498  
DEFINITION mRNA sequence.

```

VERSION      BUS01647.1  GI:22804718
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 965)
              NIH-MGC http://mgi.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgaabs-remail.nih.gov
              Tissue Procurement: ATCC
              CDNA Library Preparation: Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: L1CM2539 row: k column: 03
              High quality sequence stop: 517.
              Location/Qualifiers
                1..965
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6341498"
                /tissue_type="neuroblastoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_1lb="NIH_MGC_47"
                /note="Organ: brain; Vector: pOT7; Site_1: XhoI; Site_2:
                EcoRI; cDNA made by oligo-dt priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GGCAAGAG(G). Size-selected >500bp for average
                insert size 1.8kb. Library constructed by Ling Hong in
                the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using Zap-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.:      1,366-148      Length:      965
Score:          142.00         Matches:      142
Percent Similarity: 100.00%     Conservatve:  0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:    100.00%        Indels:       0
DB:             13            Gaps:         0

US-09-690-825-34 (1-142) x BUS01647 (1-965)
QY      1 MetGlyAlaProThrLeuProProAlaATPGInProPhleuLyAspHisArgIleSer 20
      |||
Db      54 ATGGGTGCCCGACGCTGGCCCTGCTGGAGCCCTTTCTCAAGACACCGCATCTCT 113
QY      21 ThrPhelyAsnTTPProPhleuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
      |||
Db      114 ACATTCAGAAAGTGGCCCTTCTTGAGGGCTCGGCTGCACCCCGAGGCGATGGCCAG 173
QY      41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
      |||
Db      174 GCTGGCTTATCATCGCCCTCACTGAGACGAGCCAGACTTGGCCAGAGTTCTTCTGCG 233
QY      61 PheLyGluLeuGluGlyTTPGluProAspAspAspProIleGluGlnHisLyshis 80
      |||
Db      234 TTCAGAGAGCTGGAGGCTGGGAGCGAGTGAAGCCCATAGAGAGATCAATAAAAAGCAT 293
QY      81 SerSerGlyCysAlaPheLeuSerValIleLyshGlnPheGluGluLeuThrLeuGluGlu 100
      |||
Db      294 TCATCCGGGTGGCCCTTCTTCTTGCAAGAGACACTTGAAGAAATTAACCTTGCTGAA 353
QY      101 PheLeuLyLeuAspArgGluArgAlaIysAsnLyIleAlaLyGluThrAsnLyS 120
      |||

```

```

Db      354 TTTTGAAACTGGACAGAGAAAGCCAGAAACAAATTCAGAAAGAAACCAATTAAG 413
QY      121 LyLyGluPheGluGluThrAlaIysLyValArgArgAlaIleGluGluLeuAlaIa 140
      |||
Db      414 AAGAAAGAAATTTGAGAGAACTCGAAGAAAGTGGCGGCTGCATCGAGCGCTGGCTGCC 473
QY      141 MetAsp 142
      |||
Db      474 ATCGAT 479

RESULT 59
BM463502
LOCUS
DEFINITION
5', mRNA sequence.
ACCESSION
BM463502
VERSION
BM463502.1 GI:18512544
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
              1 (bases 1 to 974)
              NIH-MGC http://mgi.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgaabs-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: L1AM2352 row: p column: 06
              High quality sequence stop: 742.
              Location/Qualifiers
                1..974
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5585909"
                /tissue_type="embryonal carcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_1lb="NIH_MGC_92"
                /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; cloned unidirectionally; oligo-dt primed.
                Average insert size 2.5 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.:      1,376-148      Length:      974
Score:          142.00         Matches:      142
Percent Similarity: 100.00%     Conservatve:  0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:    100.00%        Indels:       0
DB:             12            Gaps:         0

US-09-690-825-34 (1-142) x BM463502 (1-974)
QY      1 MetGlyAlaProThrLeuProProAlaATPGInProPhleuLyAspHisArgIleSer 20
      |||
Db      31 ATGGGTGCCCGACGCTGGCCCTGCTGGAGCCCTTTCTCAAGACACCGCATCTCT 90
QY      21 ThrPhelyAsnTTPProPhleuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
      |||
Db      91 ACATTCAGAAAGTGGCCCTTCTTGAGAGGCTGGCGCTGCACCCCGAGCGGATGGCCAG 150
QY      41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
      |||

```

```

Db      151 GCTGGCTTCACTCAGTCCCGGAGGAGCCAGACTTGGCCCGAGTTCTTCTTCG 210
Qy      61 Phelysgluuenugluyltptgiuproaspaspaspriolegluuhislyshis 80
Db      211 TTCAGGAGCTGGAAGCTGGAGCCAGATGACGACCCCATGAGGAGACATAAAAGCAT 270
Qy      81 SerSerGlyCysAlaPheleuserValIysLysGlnPhegluuhenthreunlygu 100
Db      271 TCGTCCGTTGCGCTTCTTCTTCTGCAAGAGCACTTGAAGAAATTAACCTTGAGTA 330
Qy      101 PheleuLysleuAspargluArgAlaLysAsnLysIleAlaLysGluThrasnAsnLys 120
Db      331 TTTTGAACCTGACAGAGAAAGCCAGAACCAAAATTCAGAAAGAACCAACATATAG 390
Qy      121 LysLysGluPhegluuhenthreunlyshlyValArgArgAlaIleGluGlnleuAla 140
Db      391 AAAGAAAGATTGAGGAAAGCTGAGAAAGAGTGCCTGCGCATGAGAGCTGCTGCC 450
Qy      141 Metasp 142
Db      451 ATGGAT 456

RESULT 60
BQ064557 976 bp mRNA linear EST 02-APR-2002
LOCUS AGENCOURT 6853415 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5926298
DEFINITION BQ064557.1 GI:19893301
ACCESSION BQ064557
VERSION BQ064557.1 GI:19893301
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 976)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM2099 row: 9 column: 03
High quality sequence stop: 660.
Location/Qualifiers
1..976
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5926298"
/tissue_type="lymphoma, cell line"
/lab_host="NIH_MGC_99"
/clone_lib="NIH_MGC_99"
/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-690-825-34 (1-142) x BQ064557 (1-976)
Qy      1 MetGlyAlaProthreupropoalatrnglnProPheleuLysAspHisArgIleSer 20
Db      57 ATGGGTCCCCGAGCGTTGCCCTTGAGAGGCTGGCGCTGACCCCGAGCGGATGGCCGAG 116
Qy      21 ThrPheLysAsnTrpProPheleuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db      117 ACATTCAAGAACTGGCGCTTCTTGAGAGGCTGGCGCTGACCCCGAGCGGATGGCCGAG 176
Qy      41 AlaGlyPheIleHemscysProthrgluasngluProaspLeuAlaGlnCysPhePheCys 60
Db      177 GCTGGCTTCACTCAGTCCCGGAGGAGCCAGATGACGACCCCATGAGGAGACATAAAAGCAT 236
Qy      61 Phelysgluuenugluyltptgiuproaspaspaspriolegluuhislyshis 80
Db      237 TTCAGGAGCTGGAAGCTGGAGCCAGATGACGACCCCATGAGGAGACATAAAAGCAT 296
Qy      81 SerSerGlyCysAlaPheleuserValIysLysGlnPhegluuhenthreunlygu 100
Db      297 TCGTCCGTTGCGCTTCTTCTTCTGCAAGAGCACTTGAAGAAATTAACCTTGAGTA 356
Qy      101 PheleuLysleuAspargluArgAlaLysAsnLysIleAlaLysGluThrasnAsnLys 120
Db      357 TTTTGAACCTGACAGAGAAAGCCAGAACCAAAATTCAGAAAGAACCAACATATAG 416
Qy      121 LysLysGluPhegluuhenthreunlyshlyValArgArgAlaIleGluGlnleuAla 140
Db      417 AAAGAAAGATTGAGGAAAGCTGAGAAAGAGTGCCTGCGCATGAGAGCTGCTGCC 476
Qy      141 Metasp 142
Db      477 ATGGAT 482

RESULT 61
BQ062009 985 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6459567 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5582026
DEFINITION BQ062009.1 GI:19118832
ACCESSION BQ062009
VERSION BQ062009.1 GI:19118832
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 985)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM2342 row: n column: 11
High quality sequence stop: 747.
Location/Qualifiers
1..985
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5582026"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"

```

ORIGIN

Alignment Scores:

Score: 1.37e-148

Percent Similarity: 100.00%

Length: 976

Matches: 142

Conservative: 0

/clone.lib="NIH\_MGC\_67"  
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NcoI;  
 Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.  
 Average insert size 1.75 Kb. Library constructed by Life  
 Technologies."

## ORIGIN

## Alignment Scores:

Pred. No.:	1,39e-148	Length:	985
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-09-690-825-34 (1-142) x BM80209 (1-985)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysASPHisArgIleSer 20  
 DB 63 ATGGGTGCCCCGACGCTGCCCCCTGCGACGCCCTTTCTCAAGACACCGCATCTCT

QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 123 ACATTCAAGAACTGGCCCTTCTTGAGGGGCTGCGCTTCAACCCGAGCGATGGCCGAG 182

QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 DB 183 GGTGGCTTCATCCATGCCCCCAGCTGAGAACGACGACGCTGGCCCACTGTTCTTCTGCG 242

QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
 DB 243 TTCAAGAGCTGGAGGCTGGAGCCCAATGACGACCCCATAGAGAACATTAAGAGCAT 302

QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100  
 DB 303 TCGTCGGTGGCGCTTCTTCTTCTGTCAGAAAGCATTTGAAGATTAAACCTTGTTGAA 362

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 DB 363 TTTTGAAGAACTGACAGAGAAAGAGCCAAAGAAATTGCAAGAAACCAACATATAG 422

QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140  
 DB 423 AAGAAAGATTGAGAAAGCTGCAAGAAAGTGGCGCTGCGATGAGCAGCTGCTGCC 482

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

QY 141 MetAsp 142  
 DB 483 ATGGAT 488

http://image.llnl.gov  
 Plate: LDCM2678 row: d column: 18  
 High quality sequence stop: 777.  
 Location/Qualifiers  
 1.990  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6497345"  
 /tissue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone.lib="NIH\_MGC\_101"  
 /note="Organ: lung; Vector: pMT7; Site 1: EcoRI; Site 2:  
 XhoI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC library."

## FEATURES

## Source

Alignment Scores:  
 Pred. No.: 1,39e-148 Length: 990  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

## ORIGIN

## Alignment Scores:

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysASPHisArgIleSer 20  
 DB 65 ATGGGTGCCCCGACGCTGCCCCCTGCGACGCCCTTTCTCAAGACACCGCATCTCT 124

QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 125 ACATTCAAGAACTGGCCCTTCTTGAGGGGCTGCGCTTCAACCCGAGCGATGGCCGAG 184

QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 DB 185 GGTGGCTTCATCCATGCCCCCAGCTGAGAACGACGACGCTGGCCCACTGTTCTTCTGCG 244

QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
 DB 245 TTCAAGAGCTGGAGGCTGGAGCCCAATGACGACCCCATAGAGAACATTAAGAGCAT 304

QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100  
 DB 305 TCGTCGGTGGCGCTTCTTCTTCTGTCAGAAAGCATTTGAAGATTAAACCTTGTTGAA 364

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 DB 365 TTTTGAAGAACTGACAGAGAAAGAGCCAAAGAAATTGCAAGAAACCAACATATAG 424

QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140  
 DB 425 AAGAAAGATTGAGAAAGCTGCAAGAAAGTGGCGCTGCGATGAGCAGCTGCTGCC 484

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

QY 141 MetAsp 142  
 DB 485 ATGGAT 490

ORGANISM Homo sapiens  
Mammalia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
Mammalia: Eutheria: Primates: Catarrhini, Homnidae, Homo.  
REFERENCE 1 (bases 1 to 993)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LHAM2254 row: d column: 14  
High quality sequence stop: 643.  
Location/Qualifiers  
1. 993  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5547229"  
/tissue\_type="retinoblastoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.75 kb. Library constructed by Life  
Technologies."

ORIGIN  
Alignment Scores:  
Pred. No.: 1,39e-148 Length: 993  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-690-825-34 (1-142) x BMS64739 (1-993)

QY 1 MetG1yAlaPProthreupProfoa1ATpGlnPropheluysAsph1eser 20  
Db 40 ATGGGTCGCCCGACGTTGCCCCCTGCTGCGACGCCCTTCTCAAGACCCGCACTCT 99  
QY 21 ThPhelysantPPropheluenglyCySa1aCyThProgluArgMeta1aglu 40  
Db 100 ACATTCAAGACCTGGCCCTTCTTGAGGGCTGCGCTGCACCCCGAGGATGGCCGAG 159  
QY 41 AlaGlyPhe1leHisCySPProthrgluAengluProaspLeuAlaGlnCySPhetheCyS 60  
Db 160 GCTGGCTTATCCACTGCCCCCACTGAGACGAGCACTTGGCCCACTGTTCTTCTGC 219  
QY 61 Phe1ySGluLeuGluGlyTTPGluProaPaPaPPro1leGluGluHis1yS1ySh1s 80  
Db 220 TTCAAGAGACTGAGAGGCTGGAGCCAGATGACGCCCATAGAGAACATTAAGAT 279  
QY 81 SerSerGlyCySa1aPheleuserVallySlySGlnPheGluGluLeuThre1yGlu 100  
Db 280 TCGTCCGAGTGGCTTCTTCTGTCAGAGCACTTGAAGATTAACTTGGTGA 339  
QY 101 Phe1ySGluLeuAspArg1uArg1a1ySaen1y1eal1ySGluThrAsna1y1s 120  
Db 340 TTTTCAAACTGACAGAGAAAGAGCCAGAACTTGAAGAGAAACCAACATAG 399  
QY 121 Lys1ySGluPheGluGluThra1a1yS1ySa1aArga1a1leGluGluLeuAla1a 140  
Db 400 AAGAAAGAAATTGAGAAAGCTGCAAGAAAGTCCCGTCATGAGACACTGCTGCC 459  
QY 141 Metasp 142  
|||||

Db 460 ATGCAT 465  
RESULT 64  
BM471181  
LOCUS BM471181 995 bp mRNA linear EST 05-FEB-2002  
DEFINITION AGENCOURT 6478298 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5563046  
5', mRNA sequence.  
BM471181  
ACCESSION BM471181.1 GI:18520223  
VERSION  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini, Homnidae, Homo.  
REFERENCE 1 (bases 1 to 995)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC/DCTP/DMP  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LHAM2293 row: g column: 15  
High quality sequence stop: 751.  
Location/Qualifiers  
1. 995  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5563046"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_72"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

ORIGIN  
Alignment Scores:  
Pred. No.: 1,39e-148 Length: 995  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-690-825-34 (1-142) x BM471181 (1-995)

QY 1 MetG1yAlaPProthreupProfoa1ATpGlnPropheluysAsph1eser 20  
Db 39 ATGGGTCGCCCGACGTTGCCCCCTGCTGCGACGCCCTTCTCAAGACCCGCACTCT 98  
QY 21 ThPhelysantPPropheluenglyCySa1aCyThProgluArgMeta1aglu 40  
Db 99 ACATTCAAGACCTGGCCCTTCTTGAGGGCTGCGCTGCACCCCGAGGATGGCCGAG 158  
QY 41 AlaGlyPhe1leHisCySPProthrgluAengluProaspLeuAlaGlnCySPhetheCyS 60  
Db 159 GCTGGCTTATCCACTGCCCCCACTGAGACGAGCACTTGGCCCACTGTTCTTCTGC 218  
QY 61 Phe1ySGluLeuGluGlyTTPGluProaPaPaPPro1leGluGluHis1yS1ySh1s 80  
Db 219 TTCAAGAGACTGAGAGGCTGGAGCCAGATGACGCCCATAGAGAACATTAAGAT 278  
QY 81 SerSerGlyCySa1aPheleuserVallySlySGlnPheGluGluLeuThre1yGlu 100  
Db 279 TCGTCCGAGTGGCTTCTTCTGTCAGAGCACTTGAAGATTAACTTGGTGA 338  
|||||

Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 Db 339 TTTTGAACCTGACAGAGAAAGAGCCAGCAAAATTCGCAAGAAACCAACCAATTAAG 398  
 Qy 121 LysLysGluPheGluGluThrAlaLysValArgArgAlaIleGluGluThrLeuAlaAla 140  
 Db 399 AAGAAAGAAATTGGAGAACTGCGAAGAAAGTGGCCGCGCATGAGAGCTGGCTGCC 458  
 Qy 141 MetAsp 142  
 Db 459 ATGGAT 464

RESULT 65  
 AL545180 1009 bp mRNA linear EST 31-MAY-2003  
 LOCUS  
 DEFINITION AL545180 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA  
 clone CS0D1028YH13 5-PRIME, mRNA sequence.  
 AL545180  
 VERSION AL545180.2 GI:31267016  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1009)  
 L1.W.B., Gruber,C., Jessee,J. and Polayes,D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:12877661.  
 COMMENT  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 8555.r for  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0D1028CD07Qp1&cluster=8555.r. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0D1028CD07Qp1.  
 Location/Qualifiers  
 1..1009  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1028YH13"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-cligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,41e-148 Length: 1009  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-690-825-34 (1-142) x AL545180 (1-1009)

Qy 41 AlaGlyPheIleHisCysPheProThrGluLeuGluProAspLeuAlaGlnCysPhePheCys 60  
 Db 211 GGTGGCTTCATCCTCACTGCCCCCTAGAACGAGCCAGACTGGCCAGTGTCTTCTGC 270  
 Qy 61 PheLysGluLeuGluGluYTrpGluProAspAspAspPheIleGluGluHisLysLysHis 80  
 Db 271 TTCAGAGACTGGAAGAGCTGGAGGCCAGATGACGCCCATGAGAGCAATATAAGCAT 330  
 Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluThrLeuGluGlu 100  
 Db 331 TCGTCGGGTGGCGCTTCTCTCTCTCAAGAGAGAGATTGAAGAAATTAACCTTGCTGA 390  
 Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 Db 391 TTTTGAACCTGACAGAGAAAGAGCCAGCAAAATTCGCAAGAAACCAACCAATTAAG 450  
 Qy 121 LysLysGluPheGluGluThrAlaLysValArgArgAlaIleGluGluThrLeuAlaAla 140  
 Db 451 AAGAAAGAAATTGGAGAACTGCGAAGAAAGTGGCCGCGCATGAGAGCTGGCTGCC 510  
 Qy 141 MetAsp 142  
 Db 511 ATGGAT 516

RESULT 66  
 BF033760 1031 bp mRNA linear EST 20-OCT-2000  
 LOCUS BF033760  
 DEFINITION 601454095F1 NIH\_MGC\_66 Homo sapiens CDNA clone IMAGE:3857879 5',  
 mRNA sequence.  
 BF033760  
 VERSION BF033760.1 GI:10741472  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1031)  
 NIH-MGC http://mgs.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.jnl.gov  
 Plate: LHAM9588 row: n column: 24  
 High quality sequence stop: 691.  
 Location/Qualifiers  
 1..1031  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3857879"  
 /tissue\_type="adenocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 66"  
 /note="Torgan: ovary; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.8 kb. Library constructed by Life  
 Technologies."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,44e-148 Length: 1031  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0

DB: 10 Gaps: 0

US-09-690-825-34 (1-142) x BF033760 (1-1031)

QY 1 MetGlyAlaProThrLeuProProAlaTTPGlnProPheLeuLysAspHisArgIleSer 20  
Db 7 ATGGGTCGCCCCGACGTTGCCCTTCCCTGACACCTTTCTCAAGACACCGCATCTCT 66  
QY 21 ThPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 67 ACATTCAGAGACTGGCCCTTCTTGAGAGGCTGGCTGCACCCCGAGCGATGCGCGAG 126  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
Db 127 GCTGGCTTCATCCACTGCCCCCACTGAGAGACGACACTTGGCCCAAGTTCCTCTGC 186  
QY 61 PheLeuGluLeuGluGlyTTPGluProAspAspProIleGluGluHisLysLysHis 80  
Db 187 TTCAGAGAGCTGGAAGCTGGAGGCCAGATGACGCCCATAGAGAACTATAAAGCAT 246  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
Db 247 TCGTCCGGTGGCTTCCTTCTCTGTCAGAGACAGTTTGAAGAAATTAACCTTGGTGA 306  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
Db 307 TTTTGAACCTGACAGAGAAAGAGCCAGAACAAATTGCAAGGAAACCAACATAG 366  
QY 121 LysLysGluPheGluGluThrAlaLysValArgArgAlaIleGluGluLeuAlaAla 140  
Db 367 AAGAAAGAAATTGAGGAACTGGAGAAAGTGGCCGTCATGAGAGAGCTGGCTGCC 426  
QY 141 MetAsp 142  
Db 427 ATGGAT 432

RESULT 67  
BM552163 1035 bp mRNA linear EST 20-FEB-2002  
LOCUS AGENCOURT 6543744 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5549540  
DEFINITION 5', mRNA sequence.  
ACCESSION BM552163  
VERSION BM552163.1 GI:18789795  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1035)  
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLM12260 row: d column: 21  
High quality sequence stop: 552.  
Location/Qualifiers  
1..1035  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5549540"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site\_1: Nct1;

ORIGIN

Alignment Scores:

Pred. No.: 1,44e-148 Length: 1035  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-690-825-34 (1-142) x BM552163 (1-1035)

QY 1 MetGlyAlaProThrLeuProProAlaTTPGlnProPheLeuLysAspHisArgIleSer 20  
Db 8 ATGGGTCGCCCCGACGTTGCCCTTCCCTGACACCTTTCTCAAGACACCGCATCTCT 67  
QY 21 ThPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 68 ACATTCAGAGACTGGCCCTTCTTGAGAGGCTGGCTGCACCCCGAGCGATGCGCGAG 127  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
Db 128 GCTGGCTTCATCCACTGCCCCCACTGAGAGACGACACTTGGCCCAAGTTCCTCTGC 187  
QY 61 PheLeuGluLeuGluGlyTTPGluProAspAspProIleGluGluHisLysLysHis 80  
Db 188 TTCAGAGAGCTGGAAGCTGGAGGCCAGATGACGCCCATAGAGAACTATAAAGCAT 247  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
Db 248 TCGTCCGGTGGCTTCCTTCTCTGTCAGAGACAGTTTGAAGAAATTAACCTTGGTGA 307  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
Db 308 TTTTGAACCTGACAGAGAAAGAGCCAGAACAAATTGCAAGGAAACCAACATAG 367  
QY 121 LysLysGluPheGluGluThrAlaLysValArgArgAlaIleGluGluLeuAlaAla 140  
Db 368 AAGAAAGAAATTGAGGAACTGGAGAAAGTGGCCGTCATGAGAGAGCTGGCTGCC 427  
QY 141 MetAsp 142  
Db 428 ATGGAT 433

RESULT 68  
B0048975 1047 bp mRNA linear EST 29-MAR-2002  
LOCUS AGENCOURT 6832768 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:5789750  
DEFINITION 5', mRNA sequence.  
ACCESSION B0048975  
VERSION B0048975.1 GI:19808315  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1047)  
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLM12883 row: m column: 15  
High quality sequence stop: 612.



## FEATURES

Location/Qualifiers  
1..1047

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5789750"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_92"  
/note="Organ: testis; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	145e-148	1047	142.00	142	0	0
Percent Similarity:	100.00%		100.00%	0	0	0
Best Local Similarity:	100.00%		100.00%	0	0	0
Query Match:	100.00%		100.00%	0	0	0
DB:	12					

US-09-690-825-34 (1-142) x BG048975 (1-1047)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 41 ATGGGTGCCCCGAGCTTGCCCTGCTGCGAGCCCTTTCTCAAGAGCACCCGATCTCT 100  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGln 40  
Db 101 ACATTCAAGAACTGCGCCCTCTTGAGAGGCTGCGCCCTGCACCCCGAGCGGATGCCGAG 160  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
Db 161 GCTGGCTTCATCCACTGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 220  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisLysLysHis 80  
Db 221 TTCAAGAGAGCTGGAAGCTGGAGAGCAGATGACGCCCATGAGAGAAATTAACCAT 280  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
Db 281 TCGTCCGGTTCGCTTCCTTCCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 340  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
Db 341 TTTTGAAGAACTGAG 400  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140  
Db 401 AAGAAAGAAATTTGAGAGAACTGCGAGAAAGTGGCGCTGCCATCGAGCACTGGCTGCC 460  
QY 141 MetAsp 142  
Db 461 ATGGAT 466

## RESULT 69

BG035337 1051 bp mRNA linear EST 24-JAN-2001  
LOCUS 603224964F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:4413183 5',

DEFINITION mRNA sequence.  
VERSION BG035337 GI:12429369

## KEYWORDS

## SOURCE

## ORGANISM

Human  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1051)  
NIH-MGC <http://mgi.mci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Place: LNL01038 row: p column: 16  
High quality sequence stop: 714.

## FEATURES

## source

1..1051

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4413183"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_90"  
/note="Organ: liver; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	146e-148	1051	142.00	142	0	0
Percent Similarity:	100.00%		100.00%	0	0	0
Best Local Similarity:	100.00%		100.00%	0	0	0
Query Match:	100.00%		100.00%	0	0	0
DB:	10					

US-09-690-825-34 (1-142) x BG035337 (1-1051)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 50 ATGGGTGCCCCGAGCTTGCCCTGCTGCGAGCCCTTTCTCAAGAGCACCCGATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGln 40  
Db 110 ACATTCAAGAACTGCGCCCTCTTGAGAGGCTGCGCCCTGCACCCCGAGCGGATGCCGAG 169  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
Db 170 GCTGGCTTCATCCACTGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisLysLysHis 80  
Db 230 TTCAAGAGAGCTGGAAGCTGGAGAGCAGATGACGCCCATGAGAGAAATTAACCAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
Db 290 TCGTCCGGTTCGCTTCCTTCCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
Db 350 TTTTGAAGAACTGAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140  
Db 410 AAGAAAGAAATTTGAGAGAACTGCGAGAAAGTGGCGCTGCCATCGAGCACTGGCTGCC 469  
QY 141 MetAsp 142  
Db 470 ATGGAT 475

## RESULT 70

B1518504 1052 bp mRNA linear EST 29-AUG-2001  
LOCUS B1518504

DEFINITION 603061622F1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5211005 5',  
mRNA sequence.  
ACCESSION B1518504  
VERSION B1518504.1 GI:15343296  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1052)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
plate: LLM11529 row: k column: 06  
High quality sequence stop: 834.  
Location/Qualifiers  
1..1052  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5211005"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_118"  
/note="Vector: pCMV-SPORT6, Site\_1: NotI, Site\_2: EcoRV  
(destroyed); RNA source leukocytes from anonymous pool of  
non-activated adult donors. Library is oligo-dT primed  
and directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.7 kb, insert size range  
1.2-3.3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 027. Note:  
this is a NIH\_MGC Library."

ORIGIN  
Alignment Scores:  
Pred. No.: 1..466-148 Length: 1052  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 12

US-09-690-825-34 (1-142) x B1518504 (1-1052)

QY 1 MetGlyAlaProThLeuProProAlaTrpGlnProPheLeuLysASPHisArgIleSer 20  
DB 49 ATGGGTGCCCCCAGCCCTTCCTGAGGAGGCTGCGAGCCCTTTCAGAGACACCGCATCTCT 108  
QY 21 ThrPheLysAsnTrpProPheLeuGlnGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 109 ACATTCAGAGACTGCGCCCTTCCTGAGGAGGCTGCGCCCTTCACCCGAGCGAGTGGCGAG 168  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGlnProAspLeuAlaGlnCysPhePheCys 60  
DB 169 GCTGCTCATCATCAGTCCGCCCTGAGAACAGACAGACAGCTTGGCCAGTGTCTCTTCTGCG 228  
QY 61 PheLysGluLeuGlnGlyTrpGlnProAspAspProIleGluGluHisIleLysHis 80  
DB 229 TTCAAGAGAGCTGGAAGGCTGGAGCCAGATGACGCCCATAGAGAAACATAAAACACAT 288  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 289 TCGTCCGGTGGCTTCTCTTCTCTGCAAGAGACAGTTTGAAGATTAACTTGTGTGGA 348

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
DB 349 TTTTGAACCTGCACAGAGAAAGAGCCAGACAAATTGCCAAGAAACCAACATATAG 408  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaAla 140  
DB 409 AAGAAAGAAATTTGAGGAAACTGCGAAGAAAGCGCCGTCATCAGCAGCTGGCTGCC 468  
QY 141 MetAsp 142  
DB 469 ATGCAT 474

RESULT 71  
BQ073169 1071 bp mRNA linear EST 02-APR-2002  
LOCUS AGENCOURT 6817956 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:575725  
DEFINITION 5', mRNA sequence.  
ACCESSION BQ073169  
VERSION BQ073169.1 GI:19902215  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1071)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
plate: LLM12799 row: n column: 22  
High quality sequence stop: 711.  
Location/Qualifiers  
1..1071  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:575725"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_118"  
/note="Vector: pCMV-SPORT6, Site\_1: NotI, Site\_2: EcoRV  
(destroyed); RNA source leukocytes from anonymous pool of  
non-activated adult donors. Library is oligo-dT primed  
and directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.7 kb, insert size range  
1.2-3.3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 027. Note:  
this is a NIH\_MGC Library."

ORIGIN  
Alignment Scores:  
Pred. No.: 1..486-148 Length: 1071  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 13

US-09-690-825-34 (1-142) x BQ073169 (1-1071)

QY 1 MetGlyAlaProThLeuProProAlaTrpGlnProPheLeuLysASPHisArgIleSer 20  
DB 19 ATGGGTGCCCCGACGTTGCCCCCTTCGAGAGCCCTTTCAGAGACACCGCATCTCT 78

QY 21 ThePheylsAntTppProPheleuGlUlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 79 ACATTCAAGAACTGGCCCTTCTTGGAGGCTGGCCCTCCACCCCGAGCCGATGGCCGAG 138  
 QY 41 AlaGlyPheIleHisCysPProThrgluAsnGluProAspleuAlaGlnCysPhePheCys 60  
 Db 139 GCTGGCTTCATCCACTGCCGCCACAGAGACAGACCACTGGGCCAGTGTCTTCTGCG 198  
 QY 61 PheylsGluLeuGluGlyTTrpGluProAspAspProIleGluGluHis 80  
 Db 199 TTCAAGAGCTGGAGGCTGGAGCCGAGATGACCACTTACGAGAACTTAAACCAT 258  
 QY 81 SerSerGlyCysAlaPheLeuSerVallylsGlnPheGluLeuThrLeuGlyGlu 100  
 Db 259 TCGTCCGGTTCGGCTTCTTCTGTCAAGAAAGAGTTTGAAGATTAACCTTGGTGA 318  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 Db 319 TTTTGAAGAACTGGAGAACTGCCAAGAAAGTGGCCGCTGCATCGAGCACTGGCTGCC 438  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140  
 Db 379 AAGAAAGAAATTGGAGAACTGCCAAGAAAGTGGCCGCTGCATCGAGCACTGGCTGCC 438  
 QY 141 MetAsp 142  
 Db 439 ATGGAT 444  
 RESULT 72  
 BG394560 1073 bp mRNA linear EST 12-MAR-2001  
 LOCUS 60245894F1 NIH\_MGC\_16 Homo sapiens cDNA clone IMAGE:4579367 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG394560  
 VERSION BG394560.1 GI:13288008  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1073)  
 NIH-MGC http://mgs.nci.nih.gov/.  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 Cloning strategy: Inverse Genomics, Inc.  
 Cloning distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov  
 Plate: L10M1296 row: 1 column: 24  
 High quality sequence stop: 635.  
 Location/Qualifiers  
 1..1073  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4579367"  
 /tissue\_type="retinoblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_16"  
 /note="Organ: eye; Vector: pOT37; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dt priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:  
 Pred. No.: 1,496-148 Length: 1073  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-690-825-34 (1-142) x BG394560 (1-1073)  
 QY 1 MetGlyAlaPProThleuProProAlaTppGlnProPheLeuLysAspHisArgIleSer 20  
 Db 77 ATGGGCGCCCGACAGCTGGCCCTCTCGGACGCCCTTCTCAAGAACCAACCGCATCTCT 136  
 QY 21 ThePheylsAntTppProPheleuGlUlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 137 ACATTCAAGAACTGGCCCTTCTTGGAGGCTGGCCCTCCACCCCGAGCCGATGGCCGAG 196  
 QY 41 AlaGlyPheIleHisCysPProThrgluAsnGluProAspleuAlaGlnCysPhePheCys 60  
 Db 197 GCTGGCTTCATCCACTGCCGCCACAGAGACAGACCACTGGGCCAGTGTCTTCTGCG 256  
 QY 61 PheylsGluLeuGluGlyTTrpGluProAspAspProIleGluGluHis 80  
 Db 257 TTCAAGAGCTGGAGGCTGGAGCCGAGATGACCACTTACGAGAACTTAAACCAT 316  
 QY 81 SerSerGlyCysAlaPheLeuSerVallylsGlnPheGluLeuThrLeuGlyGlu 100  
 Db 317 TCGTCCGGTTCGGCTTCTTCTGTCAAGAAAGAGTTTGAAGATTAACCTTGGTGA 376  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 Db 377 TTTTGAAGAACTGGAGAACTGCCAAGAAAGTGGCCGCTGCATCGAGCACTGGCTGCC 436  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140  
 Db 437 AAGAAAGAAATTGGAGAACTGCCAAGAAAGTGGCCGCTGCATCGAGCACTGGCTGCC 496  
 QY 141 MetAsp 142  
 Db 497 ATGGAT 502  
 RESULT 73  
 BM551159 1114 bp mRNA linear EST 20-FEB-2002  
 LOCUS AGENCOURT\_6545015 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:5738595  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BM551159  
 VERSION BM551159.1 GI:18787966  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1114)  
 NIH-MGC http://mgs.nci.nih.gov/.  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 Cloning strategy: Inverse Genomics, Inc.  
 Cloning distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: L10M12750 row: 3 column: 04  
 High quality sequence stop: 542.  
 Location/Qualifiers  
 1..1114  
 /organism="Homo sapiens"

FEATURES

source

/mol\_type="mrna"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5738595"  
 /tissue\_type="duodenal adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1ib="NIH\_MGC\_88"  
 /note="Organ: small intestine; Vector: pCMV-SPORT6;  
 Site 1: NotI; Site 2: SalI; Cloned unidirectionally;  
 oligo-dt primed. Average insert size 1.767 kb. Library  
 enriched for full-length clones and constructed by Life  
 Technologies. Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.:	1,53e-148	Length:	1114
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-09-690-825-34 (1-142) x BMS51159 (1-1114)

## ORIGIN

## Alignment Scores:

Pred. No.:	1,53e-148	Length:	1114
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-09-690-825-34 (1-142) x BQ053391 (1-1115)

## ORIGIN

## Alignment Scores:

Pred. No.:	1,54e-148	Length:	1115
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LICM2122 row: e column: 17  
 High quality sequence stop: 613.  
 Location/Qualifiers  
 1..1115  
 /organism="Homo sapiens"  
 /mol\_type="mrna"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5935096"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1ib="NIH\_MGC\_106"  
 /note="Organ: Blood; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; CDNA made by oligo-dt priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using zap-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

RESULT 74  
 BQ053391 1115 bp mRNA linear EST 29-MAR-2002  
 LOCUS BQ053391 5' mRNA sequence.  
 DEFINITION AGENCOURT.6822186 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5935096  
 ACCESSION BQ053391  
 VERSION BQ053391.1 GI:19812731  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 1115)  
 AUTHORS NIH-MGC <http://mgc.nhl.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)

RESULT 75  
 AL549362 1138 bp mRNA linear EST 31-MAY-2003  
 LOCUS AL549362

DEFINITION AL549362 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CSOD1052Y011 5-PRIME, mRNA sequence.  
ACCESSION AL549362  
VERSION AL549362.2 GI:31271180  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1138)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On Feb 15, 2001 this sequence version replaced gi:12885275.  
Contact: Genoscope - Centre National de Sequencage  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 8555.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSOD1052CE06QPl&cluster=8555.r. Contact :  
Feng Liang Email : fliang@litech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CSOD1052CE06QPl.

## FEATURES

Location/Qualifiers  
1..1138  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSOD1052Y011"  
/issue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_id="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="First strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,566-148 Length: 1138  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x AL549362 (1-1138)

QY 1 MetGlyAlaProThrLeuProProAlaTPrGlnProPheLeuLysAspHisArgTlaser 20  
DB 76 ATGGGTGCCCGACGCTGCCCCCTGCGCAGCCCTTCTCAAGACCAACCGCATCTCT 135  
QY 21 ThrPheLysAsnTPrProPheLeuGlnGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 136 ACATTCAAGAACTGGCCCTTCTTGAGGGCTGCGCTGACCCCGAGCGGATGGCCGAG 195  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
DB 196 GGTGCTTCATCATCTGCCCCCACTGAGAACGAGCAGACTGGCCCACTGTTCTTCCTGC 255  
QY 61 PheLysGluLeuGlnGlyTPrGluProAspAspProIleGluGlnHisIleLysHis 80  
DB 256 TTCAGAGAGCTGAGAGCTGGAGCAGATGACGCCCATRAGAGAACATAAAAAGCAT 315  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 316 TCGTCGGGTGGCTTCTTCTTCTGTCAGAAAGCAGTTTGAAGAAATTAACCTTGATGA 375  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
DB 376 TTTTGAAGACTGAGACAGAAAGACCAAGAACAAATTGCAAGAGAAACCAACATTAAG 435

QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaAla 140  
DB 436 AAGAAAGAAATTGAGGAAACTGCGAAGAAAGTGGCGGTCGATGAGCAGCTGGCTGCC 495  
QY 141 MetAsp 142  
DB 496 ATGGAT 501

Search completed: August 16, 2004, 03:23:12  
Job time : 2780 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_pzn model

Run on: August 16, 2004, 00:13:04 / Search time 453 Seconds

(without alignments)  
1331.665 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 142

Sequence: 1 MGAPLPPAMPOLKDHRI.....EFETAKKRAEIOLEAND 142

Scoring table:

OLIGO	Xgapop 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 3373863 segs, 212409041 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6745138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODE=frame+gzn.model -DEV=xlh  
-Q=/sgn2.1/USPRO.spool/US09690825/runat\_11082004\_141014\_13783/app\_query.fasta.1.327  
-DB=N Geneseq 29Jan04 -QMT=fastcap -SUFFIX=oligo.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=100 -DOCALLIG=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL  
-OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09690825 -CGN\_1\_1\_470/runat\_11082004\_141014\_13783 -ICPU=6  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMOUT=120 -WARN TIMOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N Geneseq 29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	142	100.0	426	2	AA231037 Open read
2	142	100.0	426	2	AA142859 Survivin
3	142	100.0	429	9	AA142859 Survivin
4	142	100.0	1024	7	AA283360 Tox1colog
5	142	100.0	1619	3	AA089303 Human sur
6	142	100.0	1619	5	AA521617 DNA encod
7	142	100.0	1619	7	ABX76199 Lung canc
8	142	100.0	1619	7	AA153734 Survivin

9	142	100.0	1619	7	AA049666	AA049666 Human sur
10	142	100.0	1619	9	ABE85228	ABE85228 Farnesyl
11	142	100.0	1662	7	ABZ58106	ABZ58106 Human sur
12	142	100.0	2404	7	ACA03988	ACA03988 CDNA down
13	142	100.0	2404	7	AAQ93052	AAQ93052 Human EPR
14	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
15	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
16	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
17	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
18	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
19	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
20	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
21	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
22	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
23	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
24	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
25	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
26	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
27	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
28	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
29	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
30	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
31	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
32	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
33	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
34	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
35	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
36	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
37	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
38	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
39	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
40	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
41	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
42	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
43	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
44	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
45	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
46	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
47	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
48	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
49	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
50	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
51	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
52	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
53	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
54	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
55	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
56	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
57	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
58	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
59	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
60	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
61	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
62	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
63	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
64	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
65	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
66	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
67	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
68	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
69	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
70	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
71	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
72	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
73	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
74	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
75	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
76	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
77	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
78	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
79	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
80	142	100.0	2404	7	ACA93052	ACA93052 Human EPR
81	142	100.0	2404	7	ACA93052	ACA93052 Human EPR

```

82 8 5.6 1734 7 ACA60388 ACA60388 Novel hum
83 8 5.6 1734 7 ACA72257 ACA72257 Novel hum
84 8 5.6 1734 7 ACD04781 ACD04781 Novel hum
85 8 5.6 1734 7 ACD18242 ACD18242 Human sec
86 8 5.6 1734 7 ACD08249 ACD08249 Human sec
87 8 5.6 1734 7 ACA88683 ACA88683 Novel hum
88 8 5.6 1734 7 ACD04997 ACD04997 Novel hum
89 8 5.6 1734 7 ACA70125 ACA70125 Human sec
90 8 5.6 1734 7 ACD12347 ACD12347 Novel hum
91 8 5.6 1734 7 ACD74262 ACD74262 Human sec
92 8 5.6 1734 7 ACD15890 ACD15890 Human sec
93 8 5.6 1734 7 ACD25458 ACD25458 Novel hum
94 8 5.6 1734 7 ACD17935 ACD17935 Human sec
95 8 5.6 1734 7 ACC88222 ACC88222 Human sec
96 8 5.6 1734 7 ACD21576 ACD21576 Human sec
97 8 5.6 1734 7 ACD18643 ACD18643 Human sec
98 8 5.6 1734 7 ACA58835 ACA58835 CDNA enco
99 8 5.6 1734 7 ABX98253 ABX98253 Human CDN
100 8 5.6 1734 7 ACD14004 ACD14004 Human PRO

```

## ALIGNMENTS

```

RESULT 1
AAZ31037
ID AAZ31037 standard; DNA; 426 BP.
XX
AC AAZ31037;
XX

```

```

DT 17-JAN-2000 (first entry)
XX

```

```

DE Open reading frame of human Survivin gene.
XX

```

```

KM Survivin; open reading frame; IAP; apoptosis inhibitor; tumour cell;
KW effector cell protease receptor-1; Epr-1; embryonic tissue; tubulin;
XX mitotic spindle; abnormal cell growth; tumorigenesis; ss.
XX

```

```

OS Homo sapiens.
XX

```

```

FH Key Location/Qualifiers
FT CDS 1..426
FT /tag= a
FT /product= "Survivin"
XX

```

```

XX WO950440-A2.
XX

```

```

PD 07-OCT-1999.
XX

```

```

PF 01-APR-1999; 99WO-US007205.
XX

```

```

PR 01-APR-1998; 98US-0080288P.
XX

```

```

XX (UYVA ) UNIV YALE.
XX

```

```

XX Altieri DC;
XX

```

```

DR WPI: 1999-59139/50.
XX

```

```

DR P-PDB; AAY49080.
XX

```

```

XX Identifying agents that modulate Survivin interactions.
XX

```

```

PS Example; Fig 8B; 56pp; English.
XX

```

```

CC This sequence is the open reading frame of the human survivin gene.
CC Survivin is a 142 amino acid protein of approximately 16.5kd. Survivin is
CC a member of the IAP family of apoptosis inhibitors, and the gene is
CC located on chromosome 17q25. The nucleic acid sequence of Survivin is
CC related to the Effector Cell Protease Receptor-1 (EPR-1). Survivin is
CC expressed in tumour cells and embryonic tissue. The sequence is used in
CC the invention which involves identifying an agent that modulates
CC interactions between Survivin and tubulin. Agents that modulate
CC interaction between Survivin and tubulin or mitotic spindles may be
CC administered to a mammal to modulate biological or pathological processes

```

```

CC mediated by Survivin, for example Survivin-mediated inhibition of
CC cellular apoptosis. Thus an inhibiting agent may be used to block
CC abnormal cell growth, for example during tumorigenesis. An agent that
CC increases interaction may be used to extend cell growth in culture
XX
SQ Sequence 426 BP; 115 A; 112 C; 113 G; 86 T; 0 U; 0 Other;

```

```

Alignment Scores:

```

```

Pred. No.: 1,15e-141 Length: 426
Score: 142.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

```

```

US-09-690-825-34 (1-142) x AAZ31037 (1-426)

```

```

QY 1 MetGlyAlaProThrLeuProProAlaATPGLnProPheLeuLysAspHisArgIleSer 20
Db 1 ATGGGTGCCCCGACGTCCTGCCCCCTGCTGCGACGCCCTTCTCAAGACACCGCATCTCT 60
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 61 ACATTCAAGACTGGCCCTTTCTTGAGAGGCTGGCCCTGCACCCCGAGCGATGGCCGAG 120
QY 41 AlaGlyPheIleHisCysEProThrgluAsngluProAspLeuAlaGlnCysPhePheCys 60
Db 121 GCTGCTTCATCCACGTCGCCCACTGAGAACGACGACGACCTTGAGCCCGACGTTCTTCTCC 180
QY 61 PheLysGluLeuGluGluGluTrpGluProAspAspAspProIleGluGluHisLysLysHis 80
Db 181 TTCAAGAGAGTGAAGAGCTGGAGGCGACGATGACGCCCATGAGAGGACATTAAGAGCAT 240
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGluGlu 100
Db 241 TCGTCGGGTGGCCCTTCTTCTGTCGACAGACAGATTGAAGATTACCTTGCTGTA 300
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrSerLys 120
Db 301 TTTTGAACATGACGACGAGAAAGCAAGACCAAAATTTGCAAGAGAAACCAACATTAAG 360
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaAla 140
Db 361 AAGAAAGATTGTAGAGAAACTGCGAAGAAAGTGGCGCGTGCATGAGACGCTGGCTGCC 420
QY 141 MetAsp 142
Db 421 ATGGAT 426

```

```

RESULT 2

```

```

AAI42859
ID AAI42859 standard; DNA; 426 BP.
XX

```

```

XX AAI42859;
XX

```

```

DT 05-AUG-2002 (first entry)
XX

```

```

DE Survivin-like protein coding sequence 6.
XX

```

```

KM Survivin-like protein; diagnosis; screening; cancer; gene; ds;
KW apoptosis abnormality; gene therapy.
XX

```

```

XX Unidentified.
XX

```

```

FH Key Location/Qualifiers
FT CDS 1..426
FT /tag= a
FT /partial=
FT /product= "Survivin-like protein 6"
FT /note= "No stop codon is given"
XX

```

```

XX W0200233071-A1.
XX

```



BD 25-APR-2002.  
 XX 16-OCT-2001; 2001WO-JP009071.  
 XX 17-OCT-2000; 2000JP-00316721.  
 XX 20-DEC-2000; 2000JP-00386809.  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX Tanaka H, Kaieda I;  
 XX WPI; 2002-435536/46.  
 DR P-PSDB; AAO14948.  
 XX Baculovirus IAP repeat domain or RING-finger domain-containing survivin-  
 PT like polypeptides and encoded DNAs, applicable in diagnosis and screening  
 PT compounds for treating various cancers and apoptosis abnormality.  
 XX Disclosure; Page 123-124; 136pp; Japanese.  
 CC The invention comprises the amino acid and coding sequences of survivin-  
 CC like proteins. The survivin-like DNA and protein sequences are useful in  
 CC diagnostics and screening compounds for treating various cancers and  
 CC apoptosis abnormality, including gene therapy. The present DNA sequence  
 CC encodes a survivin-like protein of the invention  
 XX  
 SQ Sequence 426 BP; 115 A; 112 C; 113 G; 86 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1,15e-141 Length: 426  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-690-825-34 (1-142) x AAL42859 (1-426)  
 QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 DB 1 ATGGTGCCCGGACGCTTCCCTCCGCGAGCCCTTCTCAAGAACACCGCATCTCT 60  
 QY 21 ThrPheLysAsnTrpProPheLeuGlnGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 61 ACATTCAGAACTGGCCCTTCTTGAGGGCTGCGCTGCACCCCGGACCGGATGCGCGAG 120  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 DB 121 GCTGCTTCATCCACTGCCGCCCACTGAGAACGACGACCTGGCCCGCATGTTCTTCTGC 180  
 QY 61 PheLysGluLeuGlnGlyTrpGluProAspAspProIleGluGlnHisLysLysHis 80  
 DB 181 TTCAAGAGAGCTGAGAGCTGGAGGACCATGACGACCCCATAGAGAAACATTAAGCAT 240  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 DB 241 TCGTCGGGTGGCTTCTCTTCTCTCAAGAAAGCAAGTTTGAAGATTAAACCTTGGTGA 300  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 DB 301 TTTTGAAGACTGACAGAGAAAGCAAGAAACAAATTTGCAAGAAAGAAACCAACATAG 360  
 QY 121 LysLysGluPheGlnGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
 DB 361 AAGAAAGAAATTTGAGAAATCTGCGAAGAAAGTGCCTGCGATCGAGCACTGCTGCC 420  
 QY 141 MetAsp 142  
 DB 421 ATGGAT 426  
 RESULT 3  
 ACF58071  
 ID ACF58071 standard; DNA; 429 BP.

XX ACF58071;  
 XX 15-JAN-2004 (first entry)  
 XX Human survivin protein encoding DNA.  
 DE Human survivin protein encoding DNA.  
 XX RasGAP; Aurora kinase; cytostatic; anti-tumoural; human; survivin; gene;  
 KM ds.  
 XX Homo sapiens.  
 XX OS  
 FH Key Location/Qualifiers  
 FT CDS 1..429  
 FT /tag= a  
 FT /product= "survivin"  
 XX WO2003087395-A2.  
 XX 23-OCT-2003.  
 XX 15-APR-2003; 2003WO-IB002972.  
 XX 15-APR-2002; 2002US-0372483P.  
 XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 XX  
 PI Garbay C, Gigoux V, Camonis J, L'hoste S, Samson J;  
 XX WPI; 2003-845340/78.  
 DR P-PSDB; ABRE1582.  
 XX Identifying anti-tumoral compounds, comprises determining the capacity of  
 PT a compound to inhibit interaction between RasGAP and Drosophila  
 PT melanogaster (Dm) aurora kinase, or between RasGAP, Dm aurora kinase and  
 PT aurora binding protein.  
 XX  
 XX Disclosure; Page 67; 0pp; English.  
 CC The invention relates to identifying a biologically active compound with  
 CC anti-tumoural properties, where the compound is studied for its capacity  
 CC to inhibit the interaction between: (a) RasGAP and the Drosophila  
 CC melanogaster Aurora kinase or an orthologue of the kinase, or its  
 CC fragment able to interact with RasGAP; or (b) RasGAP, an Aurora-binding  
 CC protein and the D. melanogaster Aurora kinase or an orthologue of the  
 CC kinase, or its fragment. The method is useful for identifying anti-  
 CC tumoural agents. The present sequence represents a DNA encoding a human  
 CC survivin protein, an aurora-binding protein  
 XX  
 SQ Sequence 429 BP; 116 A; 112 C; 114 G; 87 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1,15e-141 Length: 429  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-690-825-34 (1-142) x ACF58071 (1-429)  
 QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 DB 1 ATGGTGCCCGGACGCTTCCCTCCGCGAGCCCTTCTCAAGAACACCGCATCTCT 60  
 QY 21 ThrPheLysAsnTrpProPheLeuGlnGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 61 ACATTCAGAACTGGCCCTTCTTGAGGGCTGCGCTGCACCCCGGACCGGATGCGCGAG 120  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 DB 121 GCTGCTTCATCCACTGCCGCCCACTGAGAACGACGACCTGGCCCGCATGTTCTTCTGC 180

QY 61 PheLeuGluLeuGluGlyTyrGluProAspAspProIleGluGluHisLysLysHis 80  
 DB 181 TTCAGAGAGCTGGAGAGCTGGAGAGCATGACGACCCCATGAGAGAACTATMAAAGCAT 240  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 DB 241 TCGTCGGGTTGGCGCTTCTTCTGTGCAAGAGCGATTGAGAAATTAACTTGTGTGTA 300  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 DB 301 TTTTGAACCTGACAGCAAGAAAGCCCAAGAAATGCAAGAGAAACCAACATAG 360  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaIa 140  
 DB 361 AAGAAAGAAATTGAGGAAACTGCAAGAAAGTGGCGCTGATGAGAGAGCTGAGCTCC 420  
 QY 141 MetAsp 142  
 DB 421 ATGGAT 426  
 RESULT 4  
 AB283360  
 ID AB283360 standard; cDNA; 1024 BP.  
 XX  
 AC AB283360;  
 XX  
 DT 14-MAY-2003 (first entry)  
 XX  
 DE Toxicologically relevant human nucleotide sequence #519.  
 XX  
 KW Toxicologically relevant gene; toxicological response; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WC003016500-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 16-AUG-2002; 2002WC-US026514.  
 XX  
 PR 16-AUG-2001; 2001US-0313080P.  
 XX  
 PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.  
 XX  
 PI Neft RE, Dunn RT, Adkins K, Pickett GG, Kier LD, Schweiser K;  
 PI Alen P;  
 XX  
 DR MPI; 2003-268322/26.  
 XX  
 PT Determining a toxicological response to an agent, useful for screening of  
 PT drugs, comprises comparing the expression profile of one or more human  
 PT toxic response genes to a reference gene expression profile indicative of  
 PT toxicity.  
 XX  
 PS Claim 1; Page 183; 455pp; English.  
 XX  
 CC The present invention describes a method (M1) for determining a  
 CC toxicological response to an agent, which comprises comparing the  
 CC expression profile of one or more human toxic response genes to a  
 CC reference gene expression profile indicative of toxicity, and so  
 CC determining the presence of a toxic response to the agent. Also  
 CC described: (1) an array comprising one or more polynucleotides selected  
 CC from the genes corresponding to the partial sequences given in AB282842  
 CC to AB284764, or their fragments of at least 20 nucleotides, or homologues  
 CC ; and (2) determining if a gene putatively identified to be a toxic  
 CC response gene plays a role on toxic response pathways by determining the  
 CC expression profile of the gene after exposure of cells or a human subject  
 CC to a known toxic pharmaceutical or industrial agent, comprising: (a)  
 CC exposing cells to an agent or isolating cells from a human subject who  
 CC was exposed to an agent; (b) obtaining the test gene expression profile  
 CC for a putatively identified toxic response gene after exposure to a known  
 CC toxic pharmaceutical or industrial agent; and (c) comparing the test  
 CC profile to the expression profile of a gene with a similar function or

CC comparing the test profile to the expression profile of that gene after  
 CC exposure to other known toxic compounds. The methods are useful for  
 CC predicting and determining toxicological responses on a cellular, organ  
 CC or system level. The arrays comprising the human genes are useful for  
 CC toxicological screening of drugs, pharmaceutical compounds and chemicals  
 XX  
 SQ Sequence 1024 BP; 209 A; 248 C; 288 G; 279 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,62e-141 Length: 1024  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0  
 US-09-690-825-34 (1-142) x AB283360 (1-1024)  
 QY 1 MetGlyAlaProThrLeuProProAlaTyrGlnProPheLeuLysAspHisArgLieser 20  
 DB 50 ATGGGTGCCCGGACGCTGGCCCCCTGCTGGACCCCTTCTCAAGAGCAACCGCATCTT 109  
 QY 21 ThrPheLysAsnTyrProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 110 ACATTCAAGAACTGGCGCTTCTTGGAGGCGCTGGCCCTGCACCCCGAGCGGATGGCCGAG 169  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 DB 170 GCTGGCTTCAITCCACGCCCCCTGACAGACGACGACCTTGGCCAGAGCTTCTTCTTC 229  
 QY 61 PheLeuGluLeuGluGlyTyrGluProAspAspProIleGluGluHisLysLysHis 80  
 DB 230 TTCAGAGAGCTGGAGAGCTGGAGAGCATGACGACCCCATGAGAGAACTATMAAAGCAT 289  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 DB 290 TCGTCCGGTTGGCGCTTCTTCTGTCAAGAGCACTTGAAGAAATTAACTTGTGTGTA 349  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 DB 350 TTTTGAACCTGACAGCAAGAAAGCCCAAGAAATGCAAGAGAAACCAACATAG 409  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaIa 140  
 DB 410 AAGAAAGAAATTGAGGAAACTGCAAGAAAGTGGCGCTGATGAGAGAGCTGAGCTCC 469  
 QY 141 MetAsp 142  
 DB 470 ATGGAT 475  
 RESULT 5  
 AAA08903  
 ID AAA08903 standard; DNA; 1619 BP.  
 XX  
 AC AAA08903;  
 XX  
 DT 01-AUG-2000 (first entry)  
 XX  
 DE Human survivin DNA.  
 XX  
 KW Survivin; inhibitor of apoptosis; IAP; caspase inhibitor; caspase-3;  
 KW cell cycle regulation; cancer; cytostatic; antisense oligonucleotide; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 50..478  
 FT /tag= a  
 FT /product= "survivin"  
 XX  
 PN WC0200018781-A1.  
 XX  
 PD 06-APR-2000.

XX 23-SEP-1999; 99WO-US022076.  
 PF 23-SEP-1999; 99WO-US022076.  
 PR 29-SEP-1998; 98US-00163162.  
 PR 05-APR-1999; 98US-00286407.  
 XX (ISIS-) ISIS PHARM INC.  
 PA Bennett CF, Ackermann EJ, Swayze EE, Cowse LM;  
 PI WPI, 2000-293103/25.  
 DR Antisense molecules targeted to Survivin, useful for inducing apoptosis  
 XX in cancer cells.  
 PT Example 13; Page 61-62; 73pp; English.  
 PS This is the DNA sequence of human survivin (GenBank U75285), a member of  
 CC the IAP (inhibitor of apoptosis) caspase inhibitor family. Survivin has  
 CC been found to be involved in cell cycle regulation and is expressed in  
 CC the G2/M phase of the cell cycle in a cell cycle regulated manner and  
 CC associated with microtubules of the mitotic spindle. Disruption of this  
 CC interaction results in loss of survivin's anti-apoptotic function and  
 CC increased caspase-3 activity during mitosis. Caspase-3 is associated with  
 CC apoptotic cell death. It is therefore believed that survivin may  
 CC counteract a default induction of apoptosis in the G2/M phase. It is also  
 CC believed that the over expression of survivin in cancer may overcome this  
 CC apoptotic check point, allowing undesired survival and division of cancer  
 CC cells. Antisense oligonucleotides (ASO's) may be used to down regulate  
 CC endogenous survivin and to increase caspase-3-dependent apoptosis in  
 CC cells in the G2/M phase. A08910-49 are ASO's that were analyzed for  
 CC effect on survivin mRNA levels by quantitative real-time PCR. The data  
 CC obtained were averages from three experiments. It was found that ISIS  
 CC 23667 (A08925) provided 70% inhibition of Survivin and ISIS 23672  
 CC (A08930) provided 64% inhibition. The ASO's may be used to inhibit the  
 CC expression of human survivin genes for the treatment of diseases  
 CC associated with overexpression of survivin. In particular, they may be  
 CC used to induce apoptosis in tumor cells  
 XX

SO Sequence 1619 BP; 349 A; 392 C; 449 G; 429 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4,04e-141	Length:	1619
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-09-690-825-34 (1-142) x AAA08903 (1-1619)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 DB 50 ATGGATGCCCCGACCTTCCCTCCCTGCGAGCCCTTTCTCAAGAGACACCGCATCTCT 109  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 110 ACATTCAGAACTGGCCCTTCTTGGAGGGCTGCGCTGCAACCCGAGCGGATGGCCGAG 169  
 QY 41 AlaGlyPheIleHisCysProThrGlnLysAsnGluProAlaPheLeuAlaGlnCysPhePheCys 60  
 DB 170 GCTGCTCATTCATCCATGCCCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 229  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspAspProIleGluGluHisLysLysHis 80  
 DB 220 TTCAAGAGAGCTGGAGAGCTGGAGAGCATATGACGACCCCATATAGAGAAACATATAAGCAT 289  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 DB 290 TCGTCGGGTGGCTTCTCTTCTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 349  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120

DB 350 TTTTGAACCTGACAGAGAAAGCCAGAAACAAATTGCAAGAGAAACCAACATATAG 409  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGluLeuAlaIle 140  
 DB 410 AAGAAAGAAATTGAGAAACTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 469  
 QY 141 MetAsp 142  
 DB 470 ATGCAT 475

RESULT 6  
 ID AAS21617 standard; cDNA; 1619 BP.  
 XX AAS21617;  
 AC AAS21617;  
 XX 21-NOV-2001 (first entry)  
 DT DNA encoding human survivin.  
 XX Survivin; human; mouse; cytostatic; antisense oligonucleotide;  
 XX hyperproliferative condition; cancer; apoptosis; cytokinesis; ss.  
 XX Homo sapiens.  
 XX W0200157059-A1.  
 XX 09-AUG-2001.  
 PD 30-JAN-2001; 2001WO-US002939.  
 PF 02-FEB-2000; 2000US-00496694.  
 PR (ISIS-) ISIS PHARM INC.  
 XX Bennett CF, Ackermann EJ, Swayze EE, Cowse LM;  
 XX WPI, 2001-48863/53.  
 DR Novel antisense compounds for modulating the expression of Survivin and  
 XX treatment of cancer.  
 PT Example 17; Page 56; 120pp; English.  
 PS The invention relates to antisense oligonucleotides targeted to a nucleic  
 CC acid molecule encoding human Survivin, where the antisense  
 CC oligonucleotide inhibits the expression of human Survivin. These  
 CC antisense oligonucleotides are used in the treatment of an animal  
 CC suffering from a disease or condition associated with Survivin, e.g. a  
 CC hyperproliferative condition such as cancer, and comprises administering  
 CC a therapeutically or prophylactically effective amount of the antisense  
 CC oligonucleotide so that expression of Survivin is inhibited. The  
 CC oligonucleotides can also be used to treat a human suffering from a  
 CC disease or condition characterized by a reduction in apoptosis comprising  
 CC administering the antisense oligonucleotide to a human. In addition, the  
 CC antisense oligonucleotide and a cytotoxic chemotherapeutic agent e.g.  
 CC taxol or cisplatin, can be used to modulate apoptosis, cytokinesis or the  
 CC cell cycle, or inhibit the proliferation in a cancer cell by contacting  
 CC the cell with the antisense oligonucleotide. AAS21521-AAS21768 represent  
 CC Survivin nucleic acids, and antisense oligonucleotides targeted to  
 CC Survivin, used in the method of the invention  
 XX

SO Sequence 1619 BP; 349 A; 392 C; 449 G; 429 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4,04e-141	Length:	1619
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	5	Gaps:	0

US-09-690-825-34 (1-142) x AAS21617 (1-1619)

QY 1 MetGlyAlaProThleuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 50 ATGGGGCCCCACCGCTTGGCCCTGCGAGCCCTTTCTCAAGACCGCATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 110 ACATTCAAGAACTGGCCCTTCTTGAAGGCTCCGCTGCACCCCGAGACGCGATGCGGAG 169  
QY 41 AlaGlyPheIleHisCysProThrGluLeuGluProAspLeuAlaGlnCysPhePheCys 60  
DB 170 GGTGGTTTATCCTACCTGCCCCCACTGAGAACGAGCCACTGGCCAGTGTCTTCTGCG 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGlnHisLysLysHis 80  
DB 230 TTCAAGAGACTGGAAAGCTGGAGCCAGATGCGACCCCATAGAGAACATTAAGAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 290 TCGTCCGGTTGCGCTTCTTCTTCTGTCAGAAAGCATTTGAAGATTAACTTGTGTA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAspAsnLys 120  
DB 350 TTTTGAACCTGGACAGAGAAAGAGCCAGAACAAATTGCAAGGAAACCAACATTAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140  
DB 410 AAGAAAGATTGTAGAGAACTGCAAGAAAGTGCCCGCGCATGACAGCTGCTGCC 469  
QY 141 MetAsp 142  
DB 470 ATGGAT 475  
RESULT 7  
ABX76199  
ID ABX76199 standard; DNA; 1619 BP.  
AC ABX76199;  
XX  
DT 02-APR-2003 (first entry)  
XX  
DE Lung cancer-associated polynucleotide #70.  
XX  
KW Lung cancer-associated polynucleotide; gene: ds; cytostatic; emphysema;  
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
XX  
OS Unidentified.  
XX  
PN WO200286443-A2.  
XX  
PD 31-OCT-2002.  
XX  
PF 18-APR-2002; 2002WO-US012476.  
XX  
PR 18-APR-2001; 2001US-0284770P.  
PR 10-MAY-2001; 2001US-0290492P.  
PR 09-NOV-2001; 2001US-0339245P.  
PR 13-NOV-2001; 2001US-0350666P.  
PR 29-NOV-2001; 2001US-0334370P.  
PR 12-APR-2002; 2002US-0372246P.  
XX  
PA (E05B-) EOS BIOTECHNOLOGY INC.  
XX  
PI Aziz N, Murray R;  
XX  
XX  
DR WPI; 2003-093161/08.  
DR P-PDB; ABUS6477.  
XX  
PT Detecting a lung cancer-associated transcript in a cell from a patient  
for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased  
expression in lung cancer.  
XX  
PS Claim 22; Page 242; 453bp; English.  
XX  
CC The invention relates to a method for detecting a lung cancer-associated  
CC transcript in a cell from a patient, comprising contacting a biological  
CC sample from the patient with a polynucleotide that selectively hybridises  
CC to a sequence that is at least 80 % identical to a gene that exhibits  
CC increased or decreased expression in lung cancer samples. Lung cancer-  
CC associated polynucleotides and polypeptides are used for identifying a  
CC compound that modulates a lung cancer-associated polypeptide, for  
CC inhibiting proliferation of a lung cancer-associated cell to treat lung  
CC cancer in a patient and for treating a mammal having lung cancer by  
CC administering a modulatory compound identified. The methods are useful  
CC for treating lung cancer, such as small cell lung cancer, non-small cell  
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,  
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,  
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and  
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful  
CC for diagnostic purposes and as targets for screening for therapeutic  
CC compounds that modulate lung cancer, such as antibodies. Sequences  
CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the  
CC invention  
XX  
SQ Sequence 1619 BP; 349 A; 392 C; 449 G; 429 T; 0 U; 0 Other;  
XX  
Alignment Scores:  
Pred. No.: 4.04e-141 Length: 1619  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-09-690-825-34 (1-142) x ABX76199 (1-1619)  
QY 1 MetGlyAlaProThleuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 50 ATGGGGCCCCACCGCTTGGCCCTGCGAGCCCTTTCTCAAGACCGCATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
DB 110 ACATTCAAGAACTGGCCCTTCTTGAAGGCTCCGCTGCACCCCGAGACGCGATGCGGAG 169  
QY 41 AlaGlyPheIleHisCysProThrGluLeuGluProAspLeuAlaGlnCysPhePheCys 60  
DB 170 GGTGGTTTATCCTACCTGCCCCCACTGAGAACGAGCCACTGGCCAGTGTCTTCTGCG 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGlnHisLysLysHis 80  
DB 230 TTCAAGAGACTGGAAAGCTGGAGCCAGATGCGACCCCATAGAGAACATTAAGAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
DB 290 TCGTCCGGTTGCGCTTCTTCTTCTGTCAGAAAGCATTTGAAGATTAACTTGTGTA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAspAsnLys 120  
DB 350 TTTTGAACCTGGACAGAGAAAGAGCCAGAACAAATTGCAAGGAAACCAACATTAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140  
DB 410 AAGAAAGATTGTAGAGAACTGCAAGAAAGTGCCCGCGCATGACAGCTGCTGCC 469  
QY 141 MetAsp 142  
DB 470 ATGGAT 475  
RESULT 8  
AAL53734  
ID AAL53734 standard; RNA; 1619 BP.  
XX

AC AAL53734;  
 XX 07-FEB-2003 (first entry)  
 XX Survivin DNA sequence SEQ ID No 28.  
 XX Target RNA; target RNA:support-attached test compound; flow cytometry;  
 XX mass spectrometry; high-throughput screening; ds.  
 XX Homo sapiens.  
 OS WO200283837-A1.  
 XX PD 24-OCT-2002.  
 XX PF 11-APR-2002; 2002WO-US011758.  
 XX PR 11-APR-2001; 2001US-0282966P.  
 XX PA (PTCT-) PTC THERAPEUTICS INC.  
 PI Alimstead NG;  
 XX WPI; 2003-075534/07.  
 XX Identifying a test compound that binds to a target RNA molecule by  
 PT separating the detectably labeled target RNA:support-attached test  
 PT compound complex from uncomplexed target RNA molecules and test compounds  
 PT by flow cytometry.  
 XX Example; Page 81; 131pp; English.  
 XX The invention relates to a novel method for identifying a test compound  
 CC that binds to a target RNA molecule comprising separating the detectably  
 CC labeled target RNA:support-attached test compound complex from  
 CC uncomplexed target RNA molecules and test compounds. The separating  
 CC process is carried out by flow cytometry and determining a structure of  
 CC the type of test compound of the RNA:support-attached test compound  
 CC complex by mass spectrometry. The method is useful for high-throughput  
 CC screening of libraries of compounds to identify pharmaceutical leads.  
 CC This polynucleotide sequence represents a DNA sequence related to the  
 CC detecting method of the invention  
 XX SQ Sequence 1619 BP; 349 A; 392 C; 449 G; 429 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 4,04e-141 Length: 1619  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0  
 US-09-690-825-34 (1-142) x AAL53734 (1-1619)  
 QY 1 MetGlyAlaProThleuProProAlaTyrGlnProPheleuYsAspHisArgIleSer 20  
 DB 50 ATGGGTGCCCGACCTTCCCTGCTGGAGCCCTTTCACAGACACCGCATCTCT 109  
 QY 21 ThrPheIysAsnTyrProPheleuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 110 ACATTCAAGAACTGGCCCTTCTTGAGAGGCTGCGCTCAACCCGGAGCGATGGCCGAG 169  
 QY 41 AlaGlyPheIleHisCysProThleuGluAsnGluProAlaPheleuAlaGlnCysPhePheCys 60  
 DB 170 GGTGCTTCATCCATGCGCCCACTGAGAACGAGCCAGCACTGGCCCACTGTTCTTCTGCG 229  
 QY 61 PheIysGluLeuGluGlyTyrGluProAspAspProIleGluGlnHisIleYsHis 80  
 DB 230 TTCAAGAGAGCTGAAGGCTGGAGCCAGATGACGACCCCATAGAGAACATMAAAGCAT 289  
 QY 81 SerSerGlyCysAlaPheleuSerValIysGlyGlnPheGluGluLeuThleuGlyGlu 100

DB 290 TCGTCGGGTGGCTTTCCTTCTGTCAAGAACAGTGTGAAGATTAACCTTGATGA 349  
 QY 101 PheleuYsLeuAspArgGluArgAlaIysAsnIleAlaIysGluThrAspHis 120  
 DB 350 TTTTGAACCTGGACAGAGAAAGCCCAAGAACAAATTGCAGAGAAACCAACATAG 409  
 QY 121 LysIysGluPheGluGluThrAlaIysIysValArgArgAlaIleGluGlnLeuAla 140  
 DB 410 AAGAAAGATTTTGAAGAACTCCGAAGAAAGTCCCGCTGCATGAGAGAGCTGCTGCC 469  
 QY 141 MetAsp 142  
 DB 470 ATGAT 475  
 RESULT 9  
 AAD49666  
 ID AAD49666 standard; DNA; 1619 BP.  
 XX AAD49666;  
 AC 24-MAR-2003 (first entry)  
 DT Human survivin DNA.  
 XX Human survivin DNA.  
 XX Amyloidosis; haemophilia; Alzheimer's disease; atherosclerosis; cancer;  
 KW gigantism; dwarfism; hypothyroidism; hyperthyroidism; cystic fibrosis;  
 KW autoimmune disorder; aging; inflammation; diabetes; obesity; anorectic;  
 KW neurodegenerative disorder; Parkinson's disease; gene therapy; virocidic;  
 KW haemostatic; antibacterial; nootropic; neuroprotective; cytostatic; XIAP;  
 KW fungicide; human; survivin; gene; ds.  
 XX Homo sapiens.  
 OS WO200283953-A1.  
 XX PD 24-OCT-2002.  
 XX PF 11-APR-2002; 2002WO-US011757.  
 XX PR 11-APR-2001; 2001US-0282965P.  
 XX PA (PTCT-) PTC THERAPEUTICS INC.  
 PI Rando R, Welch E;  
 XX WPI; 2003-075561/07.  
 XX Identifying a test compound that binds to a target RNA molecule for  
 PT treating or preventing amyloidosis, hemophilia, cancer, gigantism,  
 PT diabetes, by contacting a detectably labeled target RNA molecule with a  
 PT library of test compounds.  
 XX Example; Page 89-90; 152pp; English.  
 XX The invention relates to a method for identifying a test compound that  
 CC binds to a target RNA molecule, which comprises contacting a detectably  
 CC labelled target RNA molecule with a library of test compounds under  
 CC conditions that permit direct binding of the labelled target RNA to a  
 CC member of the library of test compounds so that a detectably labeled  
 CC target RNA:test compound complex is formed. The method is useful for  
 CC screening libraries of compounds for those that are selectively bind to a  
 CC pre-selected target RNA. The compounds are useful for inhibiting the  
 CC formation of a specific bound RNA: host cell factor complexes in vivo.  
 CC They are also useful for treating or preventing diseases associated with  
 CC overproduction or decreased protein function, such as amyloidosis,  
 CC hemophilia, Alzheimer's disease, atherosclerosis, cancer, gigantism,  
 CC dwarfism, hypothyroidism, hyperthyroidism, autoimmune disorders, aging,  
 CC inflammation, cystic fibrosis, diabetes, obesity, neurodegenerative  
 CC disorders, Parkinson's disease or infections (bacterial, viral, fungal).  
 CC The invention is also used in gene therapy. The present sequence is human  
 CC survivin DNA. This sequence is used to illustrate the method of the  
 CC invention

SQ Sequence 1619 BP; 349 A; 392 C; 449 G; 429 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4,04e-141	Length:	1619
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	7	Gaps:	0

US-09-690-825-34 (1-142) x AAD49666 (1-1619)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20

DB 50 ATGGGGCCCCGACGCTTGCCCTCCCTGGCAGCCCTTTCTCAAGGACCGCATCTCT 109

QY 21 ThrPheLysAsnTrpProPheLeuGlyCysAlaCysThrProGluArgMetAlaGlu 40

DB 110 ACATTCAAGAACTGGCCCTCTTGAGGGCTCGGCTGCAACCCCGAGCGATGGCCAG 169

QY 41 AlaGlyPheIleHisCysProThrGluLysGluProAspLeuAlaGlnCysPhePheCys 60

DB 170 GGTGGCTTCATCCACCTGCCCCCACTAGAACGAGCCAGCTTGCCCACTGTTCTTCTGC 229

QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisIleLysIleHis 80

DB 230 TTCAAGAGCTGGAGAGCTGGAGCCAGATGACAGCCCATAGAGAACTATAAAGCAT 289

QY 81 SerSerGlyCysAlaPheLeuSerValLysIleGlnPheGluLeuThrLeuGlyGlu 100

DB 290 TCGTCCGCTTGCGCTTCTCTCTGCAAGAGCACTTGAAGAATTAACTGCTGAGAA 349

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120

DB 350 TTTTGAACCTGACAGAGAAAGCCAGCAAAATTGCAAGAGAAACCAACATAG 409

QY 121 LysLysGluPheGluGluThrAlaLysValArgArgAlaIleGluGluLeuAla 140

DB 410 AAGAAAGAAATTGAGAAACTGCGAAGAAAGTCCCGCATGACACACTGGCTCC 469

QY 141 MetAsp 142

DB 470 ATGGAT 475

RESULT 10

ID ADE85228 standard; DNA; 1619 BP.

XX ADE85228;

AC ADE85228;

XX 29-JAN-2004 (first entry)

DE Farnesyl transferase inhibitor modulated leukemia associated gene #447.

XX 88; cytostatic; farnesyl transferase inhibitor; gene expression;

OS quinolone; leukemia; cancer.

XX Homo sapiens.

XX MO2003038129-A2.

XX 08-MAY-2003.

XX 30-OCT-2002; 2002WO-US034784.

XX 30-OCT-2001; 2001US-033897P.

XX 30-OCT-2001; 2001US-0340081P.

XX 30-OCT-2001; 2001US-034038P.

XX 30-OCT-2001; 2001US-0341012P.

XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.

XX Raport M;

XX PI

XX WPI; 2003-513497/48.

DR

XX Determining whether a patient will respond to treatment with a farnesyl

PT transferase inhibitor, by analyzing the expression of gene that is

PT differentially modulated in the presence of the inhibitor.

XX Disclosure; SEQ ID NO 447; 346pp; English.

XX

XX The invention relates to a method of determining whether a patient will

CC respond to treatment with a farnesyl transferase inhibitor (FTI), by

CC analyzing the expression of gene that is differentially modulated in the

CC presence of an FTI. The method is useful for determining whether a

CC patient will respond to treatment with a FTI such as (B)-6-(amino(4-

CC chlorophenyl)-(1-methyl-1H-imidazol-5-yl)methyl)-4-(3-chlorophenyl)-1-

CC methyl-2-(1H)quinolone, monitoring the therapy of a patient, treating a

CC patient with leukemia with FTI if the analysis indicates that the patient

CC will respond. This sequence corresponds to a gene whose expression may be

CC modulated in the presence of FTI.

XX

SQ Sequence 1619 BP; 349 A; 392 C; 449 G; 429 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4,04e-141	Length:	1619
Score:	142.00	Matches:	142
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-690-825-34 (1-142) x ADE85228 (1-1619)

QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20

DB 50 ATGGGGCCCCGACGCTTGCCCTCCCTGGCAGCCCTTTCTCAAGGACCGCATCTCT 109

QY 21 ThrPheLysAsnTrpProPheLeuGlyCysAlaCysThrProGluArgMetAlaGlu 40

DB 110 ACATTCAAGAACTGGCCCTCTTGAGGGCTCGGCTGCAACCCCGAGCGATGGCCAG 169

QY 41 AlaGlyPheIleHisCysProThrGluLysGluProAspLeuAlaGlnCysPhePheCys 60

DB 170 GGTGGCTTCATCCACCTGCCCCCACTAGAACGAGCCAGCTTGCCCACTGTTCTTCTGC 229

QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisIleLysIleHis 80

DB 230 TTCAAGAGCTGGAGAGCTGGAGCCAGATGACAGCCCATAGAGAACTATAAAGCAT 289

QY 81 SerSerGlyCysAlaPheLeuSerValLysIleGlnPheGluLeuThrLeuGlyGlu 100

DB 290 TCGTCCGCTTGCGCTTCTCTCTGCAAGAGCACTTGAAGAATTAACTGCTGAGAA 349

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120

DB 350 TTTTGAACCTGACAGAGAAAGCCAGCAAAATTGCAAGAGAAACCAACATAG 409

QY 121 LysLysGluPheGluGluThrAlaLysValArgArgAlaIleGluGluLeuAla 140

DB 410 AAGAAAGAAATTGAGAAACTGCGAAGAAAGTCCCGCATGACAGACTGGCTGC 469

QY 141 MetAsp 142

DB 470 ATGGAT 475

RESULT 11

ID AB258106 standard; CDNA; 1662 BP.

XX AB258106;

AC AB258106;

XX 22-APR-2003 (first entry)

XX Human Survivin CDNA.

XX DE

XX Human; Survivin; inhibitor of apoptosis; cell death; apoptosis; cancer;  
 KM cytostatic; cardiast; neuroprotective; gene therapy; gene; ss.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 43..471  
 FT /\*tag= a  
 FT /product= "Survivin"  
 XX WO2003004606-A2.  
 PN 16-JAN-2003.  
 XX  
 PD 03-JUL-2002; 2002NC-US021002.  
 XX  
 PF 03-JUL-2001; 2001US-00898158.  
 XX  
 PR (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PA Troy CM, Shelanski ML;  
 PI WPI, 2003-210351/20.  
 XX P-PSDB; ABP72162.  
 DR  
 XX  
 PT New nucleic acid encoding an inhibitor-of-apoptosis protein, useful for  
 PT treating cancer, neurodegenerative disorder or cardiomyopathy.  
 XX  
 PS Disclosure; Fig 20B; 124bp; English.  
 XX  
 XX The present sequence is that of cDNA encoding human Survivin protein. The  
 CC invention provides a nucleic acid, such as an antisense oligonucleotide,  
 CC which specifically hybridizes to a nucleic acid encoding an inhibitor of  
 CC apoptosis protein. Survivin is an example of an inhibitor of apoptosis or  
 CC inducer of cell death protein of the invention. A claimed method for  
 CC inducing a cell's death comprises contacting the cell with the nucleic  
 CC acid under conditions permitting the nucleic acid to enter the cell;  
 CC especially the use of a vector, liposome, or a mechanical or electrical  
 CC means. The method is used to treat acute lymphocytic leukaemia, acute  
 CC myelogenous leukaemia, lung cancer, breast cancer, ovarian cancer,  
 CC prostate cancer, lymphoma, Hodgkin's disease, malignant melanoma,  
 CC neuroblastoma, renal cell carcinoma and squamous cell carcinoma (all  
 CC claimed). The invention also provides a second nucleic acid, which  
 CC specifically hybridizes to a nucleic acid encoding a protein, other than  
 CC caspase-2, that induces cell death. A claimed method for inhibiting a  
 CC cell's death comprises contacting the cell with the nucleic acid under  
 CC conditions permitting the nucleic acid to enter the cell. The method is  
 CC used to treat a neurodegenerative disorder (especially a brain disorder  
 CC or central nervous system disorder), or a heart disorder (especially  
 CC cardiomyopathy) in a human (all claimed)  
 CC  
 XX Sequence 1662 BP; 396 A; 369 C; 451 G; 426 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 4,14e-141 Length: 1662  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0  
 US-09-690-825-34 (1-142) x AB258106 (1-1662)

DB 163 GCTGCTTCATCCACTGCCACCTGAGAACGAGCCAGACTTGCCGCTTCTTCG 222  
 QY PheLysGluLeuGluGlyTTPGluProAspAspProIleGluGluHisLysHis 80  
 DB 223 TTCAAGAGCTGAGAGCTGGAGCCAGATGACACCCCAAGAGGAACATTAAGCAT 282  
 QY SerSerGlyCysAlaPheLeuSerValIleLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 DB 283 TCGTCGGGTGGCTTCTTCCTTCCTGTCAGAGAGAGATTGAAGATTAACTTGGA 342  
 QY PheLeuLysLeuAspArgGluAlaAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
 DB 343 TTTTGAACCTGAGACGAGAAAGCCAGAACAAATTGCAAGAAACCAACATTAAG 402  
 QY 121 LysLysGluPheGluGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140  
 DB 403 AAGAAAGATTGAGGAAAGTCCGAGAAAGTGGCCGCTGCATCGACACTGGCTGCC 462  
 QY 141 MetAsp 142  
 DB 463 ATGGAT 468  
 RESULT 12  
 ID ACA03988 standard; cDNA; 2404 BP.  
 AC ACA03988;  
 XX  
 DT 27-MAY-2003 (first entry)  
 XX  
 DE cDNA downregulated in senescent cells Inocyte ID NO: 251651.4.  
 XX  
 XX Human; senescence; ss; gene; cancer; proliferative disorder; leukaemia;  
 KW adenocarcinoma; lymphoma; melanoma; myeloma; sarcoma; teratocarcinoma;  
 KW adrenal gland cancer; bladder cancer; bone cancer; bone marrow cancer;  
 KW brain cancer; breast cancer; cervical cancer; colon cancer; heart cancer;  
 KW oesophageal cancer; gall bladder cancer; ganglial cancer; kidney cancer;  
 KW liver cancer; lung cancer; muscle cancer; ovarian cancer; penile cancer;  
 KW pancreatic cancer; parathyroid gland cancer; prostate cancer;  
 KW salivary gland cancer; skin cancer; small intestine cancer;  
 KW spleen cancer; stomach cancer; testicular cancer; thyroid cancer;  
 KW thyroid cancer; uterine cancer.  
 KW  
 XX Homo sapiens.  
 OS  
 XX US2002192678-A1.  
 EN 19-DEC-2002.  
 XX  
 PD 07-FEB-2002; 2002US-00071766.  
 XX  
 PF 09-FEB-2001; 2001US-0268380P.  
 PR  
 XX  
 PA (CHEN/) CHEN H.  
 XX  
 PI Chen H;  
 XX  
 DR WPI, 2003-328858/21.  
 XX  
 PT New combination comprising cDNAs or their complements, useful for  
 PT detecting changes in expression of genes encoding proteins associated  
 PT with senescence, and in diagnosing, staging or treating proliferative  
 PT diseases, e.g. cancer.  
 XX  
 XX Example 13; Page 135-136; 155bp; English.  
 PS  
 XX The invention relates to a combination comprising a plurality of cDNAs,  
 CC or their complements that are differentially expressed in cancer and  
 CC other proliferative disorders. The combination is useful in detecting  
 CC changes in expression of genes encoding proteins that are associated with  
 CC senescence and in diagnosing, staging, treating, or monitoring the  
 CC progression or treatment of subjects with proliferative diseases such as

CC cancer e.g. adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma,  
 CC sarcoma, teratocarcinoma, cancer of the adrenal gland, bladder, bone,  
 CC bone marrow, brain, breast, cervix, colon, oesophagus, gall bladder,  
 CC ganglia, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid  
 CC glands, penis, prostate, salivary glands, skin, small intestine, spleen,  
 CC stomach, testis, thymus, thyroid and uterus. The present sequence  
 CC represents cDNA of genes that are downregulated in senescent cells  
 XX

SQ Sequence 2404 BP; 525 A; 588 C; 634 G; 628 T; 0 U; 29 Other;

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 142.00	2404	142	0	0	0	0
Percent Similarity: 100.00%						
Best Local Similarity: 100.00%						
Query Match: 100.00%						

US-09-690-825-34 (1-142) x ACA03988 (1-2404)

QY 1 MetGlyAlaProThrLeuProBProAlaTrpGlnProPheLeuYsAspHisArgIleSer 20  
 DB 56 ATGGGGGCCCGACGTCGCCCCCTGCGCAGCCCTTTCAGAGACCGCATCTCT 115  
 QY 21 ThrPheYsAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 116 ACATTCAAGAACTGGCCCTCTTCTTGAGGGCTGCGCTCACCCTCGAGCGATGCGCGAG 175  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 DB 176 GCTGGCTTATCCACATGCGCCCACTGAGAAACGACGACATGGCCCAATGTTCTTCTGCG 235  
 QY 61 PheYsGluLeuGluGlyTrpGluProAspAspProIleGluGluHisIleYsHis 80  
 DB 236 TTCAGAGAGCTGGAAGGCTGGAGCGACATGACGACCCCAAGAGAAACATMAAAGCAT 295  
 QY 81 SerSerGlyCysAlaPheLeuSerValIleYsGlnPheGluGluLeuThrLeuGlyGlu 100  
 DB 296 TCGTCGGTTCGCTTCTTCTTCTCTCAAGAAAGCATTTGAAGATTACCTTGTGCGAA 355  
 QY 101 PheLeuYsLeuAspArgGluArgAlaYsAsnYsIleAlaYsGluThrAsnAsnYs 120  
 DB 356 TTTTGGAACTGGACAGAAAGACCAAGCAAAATTGCAAAAGAAACCAACAAATTAAG 415  
 QY 121 LysIysGluPheGluGluThrAlaYsYsValArgArgAlaIleGluGluLeuAla 140  
 DB 416 AAGAAAGAAATTGAGGAAATCGCAAGAAAGTGGCCCTGCAATCGAGCACTGGCTGCC 475  
 QY 141 MetAsp 142  
 DB 476 ATGGAT 481

## RESULT 13

AA093052/c  
 ID AA093052 standard; cDNA, 1165 BP.

XX AA093052;  
 AC  
 XX 25-MAR-2003 (revised)  
 DT 07-DEC-1995 (first entry)  
 XX  
 DE Human EPR-1 cDNA.  
 XX  
 KW Effector cell protease receptor-1; EPR-1; lymphocyte;  
 KW lymphoproliferative disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 37..1050  
 FT /\*tag=a  
 XX  
 PN W09520655-A1.

XX  
 PD 03-AUG-1995.  
 XX  
 PF 18-JAN-1995; 95MO-US000666.  
 XX  
 PR 28-JAN-1994; 94US-00189309.  
 XX  
 PA (SCRI) SCRIPPS RES INST.  
 XX  
 PI Altieri DC;  
 XX  
 DR WPI, 1995-275440/36.  
 DR P-PSDB; AAR6744.  
 XX  
 PT New effector cell protease receptor-1 protein - used to develop prods.  
 PT for inducing lymphocyte proliferation or diagnosis and treatment of  
 PT lymphoproliferative disorders.  
 XX  
 PS Claim 10; Page 154-155; 181pp; English.  
 XX

CC Immunoscreening of a human lymphocyte expression library in lambda. Still  
 CC using a Mab raised against human EPR-1 yielded clone lambda-104. This was  
 CC used as a probe in further screening of human cDNA libraries, and a  
 CC consensus EPR-1 cDNA sequence was deduced from 28 isolated clones. The  
 CC cDNA can be used for prodn. of recombinant EPR-1. (Updated on 25-MAR-2003  
 CC to correct PN field.)  
 XX

SQ Sequence 1165 BP; 293 A; 331 C; 300 G; 241 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 107.00	1165	138	0	1	1	0
Percent Similarity: 99.28%						
Best Local Similarity: 99.28%						
Query Match: 75.35%						

US-09-690-825-34 (1-142) x AA093052 (1-1165)

QY 1 MetGlyAlaProThrLeuProBProAlaTrpGlnProPheLeuYsAspHisArgIleSer 20  
 DB 1114 ATGGGGGCCCGACGTCGCCCCCTGCGCAGCCCTTTCAGAGACCGCATCTCT 1055  
 QY 21 ThrPheYsAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 DB 1054 ACATTCAAGAACTGGCCCTCTTCTTGAGGGCTGCGC-TGCACCCCGAGCGATGCGCGAG 996  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 DB 995 GCTGGCTTATCCACATGCGCCCACTGAGAAACGACGACATGGCCCAATGTTCTTCTGCG 936  
 QY 61 PheYsGluLeuGluGlyTrpGluProAspAspProIleGluGluHisIleYsHis 80  
 DB 935 TTCAGAGAGCTGGAAGGCTGGAGCGACATGACGACCCCAATGAGAAACATMAAAGCAT 876  
 QY 81 SerSerGlyCysAlaPheLeuSerValIleYsGlnPheGluGluLeuThrLeuGlyGlu 100  
 DB 875 TCGTCGGTTCGCTTCTTCTTCTCTCAAGAAAGCATTTGAAGATTACCTTGTGCGAA 816  
 QY 101 PheLeuYsLeuAspArgGluArgAlaYsAsnYsIleAlaYsGluThrAsnAsnYs 120  
 DB 815 TTTTGGAACTGGACAGAAAGACCAAGCAAAATTGCAAAAGAAACCAACAAATTAAG 756  
 QY 121 LysIysGluPheGluGluThrAlaYsYsValArgArgAlaIleGluGluLeuAla 139  
 DB 755 AAGAAAGAAATTGAGGAAATCGCAAGAAAGTGGCCCTGCAATCGAGCACTGGCC 699

## RESULT 14

ACA56320  
 ID ACA56320 standard; cDNA, 740 BP.

XX ACA56320;  
 AC  
 XX





## Alignment Scores:

Pred. No.: 1,72e-71 Length: 399  
 Score: 76.00 Matches: 115  
 Percent Similarity: 98.29% Conservative: 0  
 Best Local Similarity: 98.29% Mismatches: 1  
 Query Match: 53.52% Indels: 2  
 DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x AAC02317 (1-399)

QY 1 MetGlyAlaProThreuleuProProAlaTyrGlnProPheLeuLysAspHisArgIleSer 20  
 DB 52 ATGGGGCCCGACGCTGGCCCTCCCTGGCGAGCCCTTCTCAAGACACCGCCTCTCT 111  
 QY 21 ThrPheLysAsnTyrProPheLeuGluGlyCysAlaCysTyrProGluArgMetAlaGlu 40  
 DB 112 AATTTCAGACCTGGCCCTTCTTGAAGGCTCCGCTGACCCCGAGCGGATGGCCGAG 171  
 QY 41 AlaGlyPheIleHisCysProThrGluLeuGlnProAspLeuAlaGlnCysPhePheCys 60  
 DB 172 GCTGGCTTCATCCACGCGCCCACTGAGAACGACGAGACTTGGCCAGTCTTCTGCG 231  
 QY 61 PheLysGluLeuGluGlyTyrGlnProAspAspProIleGluGluHis-LysLysHis 80  
 DB 232 TTCAGAGAGCTGAGAGGCTGGAGCCAGATGACGACCCCATGAGAGASA-TAAAGACA 290  
 QY 80 sSerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGly 100  
 DB 291 TTCGTCGGGTGCGGCTTCTCTCTGTGTAAGAAAGAGTTTAAAGCTTCTGCGA 350  
 QY 100 urheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGlu 116  
 DB 351 ATTTTGAACCTGACAGAGAAAGACCAAGAAATTCGACAAAGAA 399

## RESULT 16

ADD33414 standard; DNA; 300 BP.

XX ADD33414;  
 AC ADD33414;  
 XX 15-JAN-2004 (first entry)  
 DT 15-JAN-2004 (first entry)  
 XX Mouse mitochondrial DNA sequence SEQ ID NO:1186.  
 DE Mouse mitochondrial DNA sequence SEQ ID NO:1186.  
 XX ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;  
 KM mitochondrial disease; oxidative phosphorylation dysfunction;  
 KW oxidative stress; apoptosis; aging.  
 XX Mus musculus.  
 OS Mus musculus.  
 XX MO2003020220-A2.  
 PN 13-MAR-2003.  
 PD 13-MAR-2003.  
 XX 30-AUG-2002; 2002WO-US027886.  
 PF 30-AUG-2001; 2001US-0316323P.  
 XX 31-AUG-2001; 2001CA-02356540.  
 PR (UYEM-) UNITV EMORY.  
 XX (UYEM-) UNITV EMORY.  
 PA Wallace DC, Levy S, Kerstann K, Procaccio V;  
 PI WPI; 2003-300821/29.  
 DR WPI; 2003-300821/29.  
 XX Array containing probes for genes involved in mitochondrial biology,  
 PT useful for determining mitochondrial biology gene expression profiles for  
 PT use in diagnosing pathologies and identifying biochemical pathways.  
 XX Claim 2; SEQ ID NO 1186; 201pp; English.  
 PS The invention relates to a novel array comprising at least two isolated  
 CC nucleotide molecules, each molecule having a sequence capable of uniquely

CC hybridising to a nucleic acid molecule which is an expression product of  
 CC a gene involved in mitochondrial biology. The array comprises two or more  
 CC isolated nucleic acid molecules or spots, each molecule having a sequence  
 CC chosen from sequence of 994 human probes and 2046 mouse probes. An array  
 CC of the invention is useful for determining an expression profile of a  
 CC mouse or human sample containing nucleic acid, by contacting the array  
 CC with the sample under conditions allowing selective hybridisation, and  
 CC measuring hybridisation of nucleic acid in the sample to the array to  
 CC produce an expression profile. The array is also useful for determining  
 CC an expression profile of a first labelled sample containing nucleic acid  
 CC relative to a second, differently labelled sample containing nucleic  
 CC acid. The second sample is a reference or a standard. An array is useful  
 CC for determining an expression profile diagnostic of an energy-metabolism-  
 CC related physiological condition. An array of the invention is useful for  
 CC determining mitochondrial biology gene expression profiles of organisms,  
 CC such as human, mice and closely related species, tissue and organs of  
 CC such organisms, which are useful for determining expression profiles  
 CC diagnostic of energy metabolism-related physiological conditions,  
 CC diagnosing such physiological conditions, identifying biochemical  
 CC pathways, genes, and mutations involved in such physiological conditions,  
 CC identifying therapeutic agents useful for preventing and/or treating such  
 CC physiological conditions, evaluating and/or monitoring the efficacy of  
 CC such therapies, and creating and identifying animal models of human  
 CC energy metabolism-related physiological conditions. An array is also  
 CC useful for defining expression signatures or profiles for mitochondrial  
 CC diseases, as well as distinguishing clinical disorders that result from  
 CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,  
 CC apoptosis and aging. An array of the invention contains probes of genes  
 CC not previously recognised to participate in mitochondrial biology. The  
 CC sequences shown in ADD33324-ADD33260 represent murine mitochondrial DNA  
 CC clones used to make the probes of the invention. Some sequences are not  
 CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,  
 CC 1906, 2408 and 2643.  
 XX

SQ Sequence 300 BP; 66 A; 91 C; 84 G; 59 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2.22e-34 Length: 300  
 Score: 41.00 Matches: 41  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 28.87% Indels: 0  
 DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x ADD33414 (1-300)

QY 31 CysAlaCysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrGluAsn 50  
 DB 114 TGGGCTTGACACCCAGAGCGAATGGGGAGGCTGCTCATCAGCCCTCCAGAAC 173  
 QY 51 GluProAspLeuAlaGlnCysPhePheCysPheLysGluLeuGluGlyTyrGlnProAsp 70  
 DB 174 GAGCCTGATTGGCCAGTGTTTCTTCAAGAAATGGAAGGCTGGGAACCCGAT 233  
 QY 71 Asp 71  
 DB 234 GAC 236

## RESULT 17

AAT72714 standard; DNA; 500 BP.

XX AAT72714;  
 AC AAT72714;  
 XX 16-SEP-1997 (first entry)  
 DT 16-SEP-1997 (first entry)  
 XX Mouse inhibitor of apoptosis protein homologue M1HD DNA.  
 DE Inhibitor of apoptosis protein; IAP; mammalian IAP homologue; M1HD;  
 XX degenerative disease; infectious disease; autoimmune disease; cancer;  
 KM gene therapy; diagnosis; ss.  
 XX Mus musculus.  
 OS

XX Key Location/Qualifiers  
FH CDS 48..470  
FT /\*\*tag= a  
XX  
XX MO9723501-A1.  
XX  
XX 03-JUL-1997.  
XX  
XX 20-DEC-1996; 96WO-AU000827.  
XX  
XX 22-DEC-1995; 95AU-00007275.  
XX  
XX (AMRA-) AMRAD OPERATIONS PTY LTD.  
XX  
XX Vaux DL;  
XX  
XX WPI; 1997-350966/32.  
XX  
XX P-PSDB; AAM19749.  
XX  
XX Isolated protein homologues of viral inhibitors of apoptosis - used to  
XX modulate apoptosis for treatment of degenerative, infectious or  
XX auto-immune diseases and cancer.  
XX  
XX Claim 28; Page 71-72; 136pp; English.  
XX  
XX An isolated nucleic acid molecule (AAT72714) codes for mammalian IAP  
XX homologue D (MIHD) (AAM19749), a murine homologue of baculovirus  
XX inhibitor of apoptosis protein (IAP). It was isolated following a  
XX database search for sequences showing homology to a consensus sequence  
XX (see also AAM19744) for IAP homologues. Animal IAP homologue nucleic  
XX acids (see also AAT72710-13 and AAT72715-17) can be used to produce  
XX polypeptides useful in methods for modulating apoptosis in animal cells,  
XX specifically for treatment, by inhibition, of degenerative and infectious  
XX disease or, by promotion, of cancer and autoimmune disease, and can be  
XX used for gene therapy of these diseases  
SQ Sequence 500 BP; 134 A; 131 C; 135 G; 100 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 3.6e-34 Length: 500  
Score: 41.00 Matches: 41  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 28.87% Indels: 0  
DB: 2 Gaps: 0  
US-09-690-825-34 (1-142) x AAT72714 (1-500)  
QY 31 CysAlaCysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrGluAsn 50  
Db 138 TGGCGCTGCACCCCGAGAGAAATGGCGAGGCTGCTTCATCCACTGCGCCATCCGAGAAC 197  
QY 51 GluProAspLeuAlaGlnCysPhePheCysPheIleGluLeuGluGlyTyrGluProAsp 70  
Db 198 GAGCCTGATTTGGCCAGATGTTTTCGCTTTAAGAAATGGAGAGGCTGGGAACCCGAT 257  
QY 71 Asp 71  
Db 258 GAC 260  
RESULT 18  
AAS21530  
ID AAS21530 standard; cDNA; 955 BP.  
XX  
XX AAS21530;  
XX  
XX 21-NOV-2001 (first entry)  
XX  
XX DNA encoding mouse survivin.  
XX  
XX Survivin; human; mouse; cytostatic; antisense oligonucleotide;  
XX hyperproliferative condition; cancer; apoptosis; cytokinesis; ss.  
XX

XX Mus musculus.  
OS  
XX  
XX WO200157059-A1.  
XX  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US002939.  
XX  
XX 02-FEB-2000; 2000US-00496694.  
XX  
XX (ISIS-) ISIS PHARM INC.  
XX  
XX Bennett CF, Ackermann EJ, Swayze EE, Cowseert LM;  
XX  
XX WPI; 2001-488663/53.  
XX  
XX Novel antisense compounds for modulating the expression of Survivin and  
XX treatment of cancer.  
XX  
XX Example 13; Page 80-81; 120pp; English.  
XX  
XX The invention relates to antisense oligonucleotides targeted to a nucleic  
XX acid molecule encoding human Survivin, where the antisense  
XX oligonucleotide inhibits the expression of human Survivin. These  
XX antisense oligonucleotides are used in the treatment of an animal  
XX suffering from a disease or condition associated with Survivin, e.g. a  
XX hyperproliferative condition such as cancer, and comprises administering  
XX therapeutically or prophylactically effective amount of the antisense  
XX oligonucleotide so that expression of Survivin is inhibited. The  
XX oligonucleotide can also be used to treat a human suffering from a  
XX disease or condition characterised by a reduction in apoptosis comprising  
XX administering the antisense oligonucleotide to a human. In addition, the  
XX antisense oligonucleotide and a cytotoxic chemotherapeutic agent e.g.  
XX taxol or cisplatin, can be used to modulate apoptosis, cytokinesis or the  
XX cell cycle, or inhibit the proliferation in a cancer cell by contacting  
XX the cell with the antisense oligonucleotide. AAS21521-AAS21768 represent  
XX Survivin nucleic acids, and antisense oligonucleotides targeted to  
XX Survivin, used in the method of the invention  
SQ Sequence 955 BP; 230 A; 227 C; 265 G; 233 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 6.63e-34 Length: 955  
Score: 41.00 Matches: 41  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 28.87% Indels: 0  
DB: 5 Gaps: 0  
US-09-690-825-34 (1-142) x AAS21530 (1-955)  
QY 31 CysAlaCysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrGluAsn 50  
Db 199 TGGCGCTGCACCCCGAGAGCAATGGCGAGGCTGCTTCATCCACTGCGCCATCCGAGAAC 258  
QY 51 GluProAspLeuAlaGlnCysPhePheCysPheIleGluLeuGluGlyTyrGluProAsp 70  
Db 259 GAGCCTGATTTGGCCAGATGTTTTCGCTTTAAGAAATGGAGGCTGGGAACCCGAT 318  
QY 71 Asp 71  
Db 319 GAC 321  
RESULT 19  
ABZ68313  
ID ABZ68313 standard; DNA; 121 BP.  
XX  
XX ABZ68313;  
XX  
XX 22-APR-2003 (first entry)  
XX  
XX Nucleotide sequence of a target sequence from human survivin cDNA.  
XX  
XX

XX Stem-loop; resonance energy transfer; cancer; probe;  
 KM nucleic acid detection; survival; ss.  
 XX Homo sapiens.  
 XX WO2003000933-A1.  
 XX PD 03-JAN-2003.  
 XX PF 25-JUN-2002; 2002WO-US020094.  
 XX PR 25-JUN-2001; 2001US-030672P.  
 XX PR 03-JUN-2001; 2001US-0303258P.  
 XX PA (GEOR-) GEORGIA TECH RES CORP.  
 XX PI Bao G, Tsourkas A, Xu Y;  
 XX DR WPI; 2003-210160/20.  
 XX PT Composition for detecting target nucleic acid for diagnosing cancer,  
 PT comprises dual nucleic acid probes with resonance energy transfer  
 PT moieties.  
 XX PS Example 3; Page 45; 78pp; English.  
 XX CC The specification describes a composition for the detection of a target  
 CC polynucleotide. The composition comprises two nucleic acid probes that  
 CC hybridize to a first and second nucleic acid target sequence on the  
 CC target polynucleotide. The probes form a stem-loop structure when not  
 CC bound to the target sequence, and incorporate a resonance energy transfer  
 CC donor or acceptor moiety. When both probes hybridize to the target, an  
 CC observable and detectable interaction occurs. The method is useful for  
 CC detecting a subject polynucleotide in vivo. The probes are useful for  
 CC detecting nucleic acid target hybridization, for the identification and  
 CC quantification of tissue and cell-specific gene expression levels,  
 CC including response to external stimuli, such as drug candidates, and  
 CC genetic variations associated with disease such as cancer. The present  
 CC sequence represents a fragment of human survival CDNA. Probes AB26314-15  
 CC hybridize to this sequence, and are used to demonstrate the method of the  
 CC invention.  
 XX SQ Sequence 121 BP; 19 A; 44 C; 35 G; 23 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1,09e-33 Length: 121  
 Score: 40.00 Matches: 40  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 28.17% Indels: 0  
 DB: Gaps: 0  
 US-09-690-825-34 (1-142) x AB268313 (1-121)  
 QY 1 MetGlyAlaProTherLeuProProAlaTrpGlnProPhenLeuLysAspH;ArgIleSer 20  
 Db 1 AAGGGAGGCCCGACCTTGGCCCTGCGAGCGCCCTTTCACAGACACCGCATCTCT 60  
 QY 21 ThrPheLysAsnTrpProPhenLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 61 ACATTCAAGAACTGCGCCCTTCTTGAGAGGGCTGCGCTGCACCCGAGCGGATGCGCGAG 120  
 RESULT 20  
 ADB67958/C  
 ID ADB67958 standard; CDNA; 1050 BP.  
 XX ADB67958;  
 XX 04-DEC-2003 (first entry)  
 XX Human lung specific nucleic acid sequence SEQ ID NO:57.

XX human; lung specific polypeptide; LSP; lung specific nucleic acid; LSNA;  
 KM lung; cytosolic; vaccine; gene therapy; immune response; lung cancer;  
 KM metastasis; gene; ss.  
 XX Homo sapiens.  
 XX WO2003020899-A2.  
 XX PD 13-MAR-2003.  
 XX PF 29-AUG-2002; 2002WO-US027771.  
 XX PR 31-AUG-2001; 2001US-0316260P.  
 XX PA (DIAD-) DIADEXUS INC.  
 XX PI Sun Y, Liu C, Chen S;  
 XX DR WPI; 2003-300880/29.  
 XX DR P-FSDB; ADB67863.  
 XX PT Novel lung specific polypeptides and nucleic acids, useful for  
 PT identifying, diagnosing, monitoring, staging, imaging and treating lung  
 PT cancer and non-cancerous disease states in lung.  
 XX PS Example 1; Page 200-201; 263pp; English.  
 XX CC The present invention describes human lung specific polypeptides (LSP,  
 CC (I)) and lung specific nucleic acids (LSNA, (II)). Also described: (1)  
 CC determining (M1) the presence of an LSNA in a sample; (2) a vector (II)  
 CC comprising (II); (3) a host cell comprising (II); (4) a polypeptide (IV)  
 CC encoded by (II); (5) production of (IV); (6) an antibody (V) or its  
 CC fragment that specifically binds to an LSP; (7) determining (M2) the  
 CC presence of an LSP in a sample; (8) a kit for detecting a risk of cancer  
 CC or presence of cancer in a patient, comprising a unit for determining the  
 CC presence of (I) or (II) in a sample of the patient; and (9) a vaccine  
 CC comprising (I) or (II). LSP and LSNA sequences have cytosolic  
 CC activities, and can be used in vaccines, gene therapy, and as inducers of  
 CC an immune response. (I) and (II) can be used for diagnosing or monitoring  
 CC the presence and metastases of lung cancer in a patient, by determining  
 CC the amount of (I) or (II) in a sample of a patient, and comparing the  
 CC amount of the determined nucleic acid molecule or the polypeptide in the  
 CC sample of the patient to the amount of the lung specific marker in a  
 CC normal control, where the difference in the amount of the nucleic acid  
 CC molecule or the polypeptide in the sample compared to the amount of the  
 CC nucleic acid molecule or the polypeptide in the normal control is  
 CC associated with the presence of lung cancer. (I) and (V) are useful for  
 CC treating a patient with lung cancer, by administering (I) or (V), where  
 CC the administration of (I) or (V) induces an immune response against the  
 CC lung cancer cell expressing the nucleic acid molecule or a polypeptide.  
 CC (I), (II) and (V) are also useful for identifying, diagnosing,  
 CC monitoring, staging, imaging and treating lung cancer and non-cancerous  
 CC disease states in lung, for identifying lung tissue, and monitoring,  
 CC identifying and/or designing agonists and antagonists of (I). (II) can be  
 CC used in gene therapy, for producing transgenic animals and cells, and for  
 CC producing engineered lung tissue for treatment and research. (II) is also  
 CC useful for detecting or amplifying nucleic acid molecules that have  
 CC similar or identical nucleic acid sequences compared to (II). (I) is  
 CC useful for producing engineered lung tissue. (V) is useful for  
 CC identifying (I). (II) is also useful for driving in vivo expression of  
 CC (I). The present sequence represents a human LSNA from the present  
 CC invention.  
 XX SQ Sequence 1050 BP; 316 A; 231 C; 178 G; 325 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 9.72e-32 Length: 1050  
 Score: 39.00 Matches: 39  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 27.46% Indels: 0  
 DB: Gaps: 0



CC The invention provides antisense compounds capable of inhibiting the  
CC expression of human Her-3, a member of epidermal growth factor (EGF)  
CC family of receptor/tyrosine kinases. The antisense oligonucleotides are  
CC useful for inhibiting the expression of Her-3 in cells or tissues. They  
CC are commonly used as research reagents and in diagnostics for example, to  
CC elucidate the function of particular genes. The antisense compounds are  
CC also useful for distinguishing between functions of various members of a  
CC biological pathway and for research use. They are also utilized for  
CC diagnostics, therapeutics, prophylaxis and in kits. They are useful  
CC for prophylactically, e.g. to prevent or delay infection, inflammation or  
CC tumor formation. The present sequence represents the human Her-3 DNA  
CC sequence (GenBank Accession No. U75285)  
XX

Sequence 14796 BP; 3482 A; 3488 C; 3888 G; 3938 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:
1,18e-30	14796	39
Score:	39.00	Conservative: 0
Percent Similarity:	100.00%	Mismatches: 0
Best Local Similarity:	100.00%	Indels: 0
Query Match:	27.46%	Gaps: 0

DB: 4

US-09-690-825-34 (1-142) x AAG47531 (1-14796)

QY 75 GluGluHisIysLysHisSerSerGlyCysAlaPheLeuSerValIysLysGlnPheGlu 94  
DB 5159 GAGGAACATATAAAGCATTCGTCGCGTTCCTTCTGTCAAGAGCAGTTTGA 5218

QY 95 GluLeuThrLeuGlyGlnPheLeuLysLeuAspArgGluArgAlaLysAsnIle 113  
DB 5219 GAATTAACTCTTGTAATTTTGAACCTGACAGAGAAAGCCAGAAACAAAT 5275

RESULT 23  
AAS21523  
ID AAS21523 standard; cDNA; 14796 BP.  
XX  
AC AAS21523;  
XX  
DT 21-NOV-2001 (first entry)  
XX  
DE DNA encoding human survivin.  
XX  
KW Survivin; human; mouse; cytostatic; antisense oligonucleotide;  
KM hyperproliferative condition; cancer; apoptosis; cytokinesis; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157059-A1.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US002939.  
XX  
PR 02-FEB-2000; 2000US-00496694.  
XX  
PA (ISIS-) ISIS PHARM INC.  
XX  
PI Bennett CF, Ackermann EJ, Swayze EF, Cowsett LM;  
XX  
DR WPI, 2001-488663/53.  
XX  
XX Novel antisense compounds for modulating the expression of Survivin and  
XX treatment of cancer.  
XX  
PS Example 13; Page 74-79; 120pp; English.  
XX  
CC The invention relates to antisense oligonucleotides targeted to a nucleic  
CC acid molecule encoding human Survivin, where the antisense  
CC oligonucleotide inhibits the expression of human Survivin. These  
CC antisense oligonucleotides are used in the treatment of an animal  
CC suffering from a disease or condition associated with Survivin, e.g. a  
CC hyperproliferative condition such as cancer, and comprises administering

CC a therapeutically or prophylactically effective amount of the antisense  
CC oligonucleotide so that expression of Survivin is inhibited. The  
CC oligonucleotides can also be used to treat a human suffering from a  
CC disease or condition characterised by a reduction in apoptosis comprising  
CC administering the antisense oligonucleotide to a human. In addition, the  
CC antisense oligonucleotide and a cytotoxic chemotherapeutic agent e.g.  
CC taxol or cisplatin, can be used to modulate apoptosis, cytokinesis or the  
CC cell cycle, or inhibit the proliferation in a cancer cell by contacting  
CC the cell with the antisense oligonucleotide. AAS2151-AAS21768 represent  
CC Survivin nucleic acids, and antisense oligonucleotides targeted to  
CC Survivin, used in the method of the invention  
XX

Sequence 14796 BP; 3482 A; 3488 C; 3888 G; 3938 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:
1,18e-30	14796	39
Score:	39.00	Conservative: 0
Percent Similarity:	100.00%	Mismatches: 0
Best Local Similarity:	100.00%	Indels: 0
Query Match:	27.46%	Gaps: 0

DB: 5

US-09-690-825-34 (1-142) x AAS21523 (1-14796)

QY 75 GluGluHisIysLysHisSerSerGlyCysAlaPheLeuSerValIysLysGlnPheGlu 94  
DB 5159 GAGGAACATATAAAGCATTCGTCGCGTTCCTTCTGTCAAGAGCAGTTTGA 5218

QY 95 GluLeuThrLeuGlyGlnPheLeuLysLeuAspArgGluArgAlaLysAsnIle 113  
DB 5219 GAATTAACTCTTGTAATTTTGAACCTGACAGAGAAAGCCAGAAACAAAT 5275

RESULT 24  
ABL68655  
ID ABL68655 standard; DNA; 14796 BP.  
XX  
AC ABL68655;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Kidney cancer related gene sequence SEQ ID NO:6992.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KM stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US010838.  
XX  
PR 05-JUN-2000; 2000US-0209473P.  
PR 05-JUN-2000; 2000US-0209531P.  
PR 18-SEP-2000; 2000US-0233133P.  
PR 18-SEP-2000; 2000US-0233617P.  
PR 20-SEP-2000; 2000US-0234009P.  
PR 20-SEP-2000; 2000US-0234034P.  
PR 22-SEP-2000; 2000US-0234509P.  
PR 22-SEP-2000; 2000US-0234567P.  
PR 25-SEP-2000; 2000US-0234923P.  
PR 25-SEP-2000; 2000US-0234924P.  
PR 25-SEP-2000; 2000US-0235077P.  
PR 25-SEP-2000; 2000US-0235082P.  
PR 25-SEP-2000; 2000US-0235134P.  
PR 25-SEP-2000; 2000US-0235280P.  
PR 26-SEP-2000; 2000US-0235637P.  
PR 26-SEP-2000; 2000US-0235638P.  
PR 27-SEP-2000; 2000US-0235711P.

27-SEP-2000; 2000US-0235720P.  
PR 27-SEP-2000; 2000US-0235840P.  
PR 27-SEP-2000; 2000US-0235863P.  
PR 28-SEP-2000; 2000US-0236028P.  
PR 28-SEP-2000; 2000US-0236032P.  
PR 28-SEP-2000; 2000US-0236033P.  
PR 28-SEP-2000; 2000US-0236034P.  
PR 28-SEP-2000; 2000US-0236109P.  
PR 28-SEP-2000; 2000US-0236111P.  
PR 29-SEP-2000; 2000US-0236842P.  
PR 29-SEP-2000; 2000US-0236891P.  
PR 02-OCT-2000; 2000US-0237172P.  
PR 02-OCT-2000; 2000US-0237173P.  
PR 02-OCT-2000; 2000US-0237294P.  
PR 02-OCT-2000; 2000US-0237295P.  
PR 02-OCT-2000; 2000US-0237316P.  
PR 03-OCT-2000; 2000US-0237425P.  
PR 03-OCT-2000; 2000US-0237598P.  
PR 03-OCT-2000; 2000US-0237604P.  
PR 03-OCT-2000; 2000US-0237606P.  
PR 03-OCT-2000; 2000US-0237608P.  
PR 01-NOV-2000; 2000US-0244867P.  
PR 01-NOV-2000; 2000US-0245084P.  
PA (AVAL-) AVALON PHARM.  
XX  
XX  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppe DR, Weaver Z;  
XX  
XX WPI; 2002-188264/24.  
XX  
XX  
XX Screening for anti-neoplastic agent involves exposing cells to a chemical  
PT agent to be tested for anti-neoplastic activity, and determining a change  
PT in expression of a gene of a signature gene set.  
XX  
XX  
PS Claim 1; SEQ ID NO 6992; 44pp; English.  
XX  
XX The present invention describes a method (M1) for screening for an anti-  
CC neoplastic agent. The method involves exposing cells to a chemical agent  
CC to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 847 sequences (given in ABL66364  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening an  
CC anti-neoplastic agent, and can be used for producing a product which is  
CC the data collected with respect to the anti-neoplastic agent as a result  
CC of M1, and the data is sufficient to convey the chemical structure and/or  
CC properties of the agent. M1 can be used in the treatment of cancer such  
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms  
CC tumour  
XX  
XX Sequence 14796 BP; 3482 A; 3488 C; 3888 G; 3938 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 1.18e-30 Length: 14796  
Score: 39.00 Matches: 39  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 27.46% Indels: 0  
DB: 6 Gaps: 0  
US-09-690-825-34 (1-142) x ABL66655 (1-14796)  
QY 75 GIUGIUIHSLYSLYSHISSErserGIYCSAlaPhelneuserVAllySLYGLnPhesglu 94  
Db 5159 GAGGAACATATAAACATTCGCGGTTCCGTTTCCTTCTGTCAAGAACCATTTGAA 5218  
QY 95 GIULeuThrLeuGIYGLuPhelneulysleuSparGIuArGIAluYAsnlySalle 113

Db 5219 GAATTACCCCTGGTGAATTTTGGAACTGGACAGAGAAAGCCAGAACAAATT 5275  
RESULT 25  
ABL66326  
ID ABL66326 standard; DNA, 14796 BP.  
XX  
XX  
XX ABL66326;  
AC  
XX  
XX 15-MAY-2002 (first entry)  
DT  
XX  
XX Lung cancer related gene sequence SEQ ID NO:4663.  
DE  
XX  
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;  
KW gene; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200194629-A2.  
PN  
XX  
XX 13-DEC-2001.  
PD  
XX  
XX 30-MAY-2001; 2001WO-US010838.  
PF  
XX  
XX 05-JUN-2000; 2000US-0209473P.  
PR 05-JUN-2000; 2000US-0209531P.  
PR 18-SEP-2000; 2000US-0231133P.  
PR 18-SEP-2000; 2000US-0233617P.  
PR 20-SEP-2000; 2000US-0234009P.  
PR 20-SEP-2000; 2000US-0234034P.  
PR 20-SEP-2000; 2000US-0234052P.  
PR 22-SEP-2000; 2000US-0234509P.  
PR 22-SEP-2000; 2000US-0234567P.  
PR 25-SEP-2000; 2000US-0234923P.  
PR 25-SEP-2000; 2000US-0234924P.  
PR 25-SEP-2000; 2000US-0235077P.  
PR 25-SEP-2000; 2000US-0235082P.  
PR 25-SEP-2000; 2000US-0235134P.  
PR 25-SEP-2000; 2000US-0235280P.  
PR 26-SEP-2000; 2000US-0235637P.  
PR 26-SEP-2000; 2000US-0235638P.  
PR 27-SEP-2000; 2000US-0235711P.  
PR 27-SEP-2000; 2000US-0235720P.  
PR 27-SEP-2000; 2000US-0235840P.  
PR 27-SEP-2000; 2000US-0235863P.  
PR 28-SEP-2000; 2000US-0236028P.  
PR 28-SEP-2000; 2000US-0236032P.  
PR 28-SEP-2000; 2000US-0236033P.  
PR 28-SEP-2000; 2000US-0236034P.  
PR 28-SEP-2000; 2000US-0236109P.  
PR 28-SEP-2000; 2000US-0236111P.  
PR 28-SEP-2000; 2000US-0236842P.  
PR 28-SEP-2000; 2000US-0236891P.  
PR 29-SEP-2000; 2000US-0237172P.  
PR 02-OCT-2000; 2000US-0237173P.  
PR 02-OCT-2000; 2000US-0237172P.  
PR 02-OCT-2000; 2000US-0237294P.  
PR 02-OCT-2000; 2000US-0237295P.  
PR 02-OCT-2000; 2000US-0237316P.  
PR 03-OCT-2000; 2000US-0237425P.  
PR 03-OCT-2000; 2000US-0237598P.  
PR 03-OCT-2000; 2000US-0237604P.  
PR 03-OCT-2000; 2000US-0237606P.  
PR 03-OCT-2000; 2000US-0237608P.  
PR 01-NOV-2000; 2000US-0244867P.  
PR 01-NOV-2000; 2000US-0245084P.  
XX  
XX (AVAL-) AVALON PHARM.  
PA  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppe DR, Weaver Z;

XX WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a chemical

PT agent to be tested for anti-neoplastic activity, and determining a change

PT in expression of a gene of a signature gene set.

XX Claim 1; SEQ ID NO 4663; 44pp; English.

XX The present invention describes a method (M1) for screening for an anti-

CC neoplastic agent. The method involves exposing cells to a chemical agent

CC to be tested for anti-neoplastic activity, determining a change in

CC expression of at least one gene (I) of a signature gene set, where (I)

CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664

CC to ABL70110), or is at least 95% identical to (S), where a change in

CC expression is indicative of anti-neoplastic activity. (I) has cytostatic

CC activity and can be used in gene therapy. M1 can be used for screening an

CC anti-neoplastic agent, and can be used for producing a product which is

CC the data collected with respect to the anti-neoplastic agent as a result

CC of M1, and the data is sufficient to convey the chemical structure and/or

CC properties of the agent. M1 can be used in the treatment of cancer such

CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,

CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell

CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous

CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms

CC tumour

XX

XX Sequence 14796 BP; 3482 A; 3488 C; 3888 G; 3938 T; 0 U; 0 Other;

XX

Alignment Scores:

Pred. No.: 1,186-30 Length: 14796

Score: 39.00 Matches: 39

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 27.46% Indels: 0

DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x ABL66326 (1-14796)

QY 75 GUGUHHHSLYSLEHSHSERSERGLYCYSALAPHELEUSERVALLYSLEGLPHGCU 94

DB 5159 GAGGAAACATTAACCAATTCGTCGCGTTCCTTCCTTCGTCMAAGACAGTTCGA 5218

QY 95 GULGUTHPLGUGLGLUPHEULYSLEUSPAAGUAGUAGALYSASNLVLE 113

DB 5219 GAATTACCTCTTGATTTTGAACCTGACAGAGAAAGCCAGAACAAATT 5215

RESULT 26

ABL65663

ID ABL65663 standard; DNA; 14796 BP.

XX

XX ABL65663;

XX

XX 15-MAY-2002 (first entry)

XX

XX Lung cancer related gene sequence SEQ ID NO:4000.

XX

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

XX cytostatic; gene therapy; anti-neoplastic; Wilms tumour; adenocarcinoma;

XX gene; de.

XX

XX Homo sapiens.

XX

XX WO200194629-A2.

XX

XX 13-DEC-2001.

XX

XX 30-MAY-2001; 2001WO-US010838.

XX

XX 05-JUN-2000; 2000US-02039473P.

XX

XX 05-JUN-2000; 2000US-02039531P.

XX

XX 18-SEP-2000; 2000US-0233133P.

PR 18-SEP-2000; 2000US-0233617P.

PR 20-SEP-2000; 2000US-0234009P.

PR 20-SEP-2000; 2000US-0234034P.

PR 20-SEP-2000; 2000US-0234052P.

PR 22-SEP-2000; 2000US-0234509P.

PR 22-SEP-2000; 2000US-0234567P.

PR 25-SEP-2000; 2000US-0234923P.

PR 25-SEP-2000; 2000US-0234924P.

PR 25-SEP-2000; 2000US-0235077P.

PR 25-SEP-2000; 2000US-0235082P.

PR 25-SEP-2000; 2000US-0235134P.

PR 25-SEP-2000; 2000US-0235280P.

PR 25-SEP-2000; 2000US-0235637P.

PR 26-SEP-2000; 2000US-0235638P.

PR 27-SEP-2000; 2000US-0235711P.

PR 27-SEP-2000; 2000US-0235720P.

PR 27-SEP-2000; 2000US-0235840P.

PR 27-SEP-2000; 2000US-0235843P.

PR 28-SEP-2000; 2000US-0235828P.

PR 28-SEP-2000; 2000US-0236032P.

PR 28-SEP-2000; 2000US-0236033P.

PR 28-SEP-2000; 2000US-0236034P.

PR 28-SEP-2000; 2000US-0236109P.

PR 28-SEP-2000; 2000US-0236111P.

PR 28-SEP-2000; 2000US-0236482P.

PR 29-SEP-2000; 2000US-0236891P.

PR 02-OCT-2000; 2000US-0237172P.

PR 02-OCT-2000; 2000US-0237173P.

PR 02-OCT-2000; 2000US-0237278P.

PR 02-OCT-2000; 2000US-0237294P.

PR 02-OCT-2000; 2000US-0237295P.

PR 02-OCT-2000; 2000US-0237316P.

PR 03-OCT-2000; 2000US-0237425P.

PR 03-OCT-2000; 2000US-0237598P.

PR 03-OCT-2000; 2000US-0237604P.

PR 03-OCT-2000; 2000US-0237606P.

PR 03-OCT-2000; 2000US-0237608P.

PR 01-NOV-2000; 2000US-0244867P.

PR 01-NOV-2000; 2000US-0245084P.

XX

XX (AVAL-) AVALON PHARM.

XX

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

PI Seppert DR, Weaver Z;

XX

XX WPI; 2002-188264/24.

XX

XX Screening for anti-neoplastic agent involves exposing cells to a chemical

PT agent to be tested for anti-neoplastic activity, and determining a change

PT in expression of a gene of a signature gene set.

XX

XX Claim 1; SEQ ID NO 4000; 44pp; English.

XX

XX The present invention describes a method (M1) for screening for an anti-

CC neoplastic agent. The method involves exposing cells to a chemical agent

CC to be tested for anti-neoplastic activity, determining a change in

CC expression of at least one gene (I) of a signature gene set, where (I)

CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664

CC to ABL70110), or is at least 95% identical to (S), where a change in

CC expression is indicative of anti-neoplastic activity. (I) has cytostatic

CC activity and can be used in gene therapy. M1 can be used for screening an

CC anti-neoplastic agent, and can be used for producing a product which is

CC the data collected with respect to the anti-neoplastic agent as a result

CC of M1, and the data is sufficient to convey the chemical structure and/or

CC properties of the agent. M1 can be used in the treatment of cancer such

CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,

CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell

CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous

CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms

CC tumour

XX

XX Sequence 14796 BP; 3482 A; 3488 C; 3888 G; 3938 T; 0 U; 0 Other;



## Alignment Scores:

Pred. No.: 1.18e-30 Length: 14796  
 Score: 39.00 Matches: 39  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 27.46% Indels: 0  
 DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x ABL65663 (1-14796)

Qy 75 GluGluHisLysLysHisSerSerGlyCysAlaPheLeuSerValLysLysGlnPheGlu 94  
 Db 5159 GAGGAACATMAAAAGCATTCGTCGCTTCCTTCCTTCGTCAGAGGACAGTTTGA 5218

Qy 95 GluLeuThrLeuGlyGlnPheLeuLysLeuAspArgGluArgAlaLysAsnLysIle 113  
 Db 5219 GAATTAACTTGGATTTTGAATGAGACAGAGAGGACCAAGACAAATTT 5275

## RESULT 27

ABN96924  
 ID ABN96924 standard; DNA; 14796 BP.

AC ABN96924;

DT 13-AUG-2002 (first entry)

DE Gene #3422 used to diagnose liver cancer.

XX Gene, liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
 KM metastatic liver tumour; cytostatic; expression profile; disease state;  
 KM disease progression; drug toxicity; drug efficacy; drug metabolism.

XX Homo sapiens.

OS WO200229103-A2.

PN 11-APR-2002;

PD 02-OCT-2001; 2001WO-US030589.

PF 02-OCT-2000; 2000US-0237054P.

PR (GENE-) GENE LOGIC INC.

PA Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

PI WPI; 2002-426119/45.

DR Diagnosing and detecting the progression of liver cancer, hepatocellular  
 XX carcinoma or metastatic liver tumor in a patient, involves detecting the  
 PT level of expression of two or more genes in a liver tissue sample.

PS Claim 1; SEQ ID NO 3422; 298bp; English.

XX The invention relates to a novel method for diagnosing and detecting the  
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
 CC tumor in a patient, and differentiating metastatic liver cancer from  
 CC hepatocellular carcinoma in a patient, involving detecting the level of  
 CC expression of two or more genes represented in ABN93503-ABN97455 in a  
 CC tissue sample. The method of the invention has hepatotropic, and  
 CC cytostatic activity. The method is useful for diagnosing and detecting  
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
 CC liver carcinoma in a patient. The method is useful for identifying  
 CC expression profiles which serve as useful diagnostic markers as well as  
 CC markers that can be used to monitor disease states, disease progression,  
 CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 14796 BP; 3482 A; 3488 C; 3888 G; 3938 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.18e-30 Length: 14796

Score: 39.00 Matches: 39  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 27.46% Indels: 0  
 DB: 6 Gaps: 0

US-09-690-825-34 (1-142) x ABN96924 (1-14796)

Qy 75 GluGluHisLysLysHisSerSerGlyCysAlaPheLeuSerValLysLysGlnPheGlu 94  
 Db 5159 GAGGAACATMAAAAGCATTCGTCGCTTCCTTCCTTCGTCAGAGGACAGTTTGA 5218

Qy 95 GluLeuThrLeuGlyGlnPheLeuLysLeuAspArgGluArgAlaLysAsnLysIle 113  
 Db 5219 GAATTAACTTGGATTTTGAATGAGACAGAGAGGACCAAGACAAATTT 5275

## RESULT 28

AAZ31036  
 ID AAZ31036 standard; DNA; 403 BP.

AC AAZ31036;

DT 17-JAN-2000 (first entry)

DE 5' flanking region of the human Survivin gene.

XX Survivin; 5' flanking region; IAP; apoptosis inhibitor; tumour cell;  
 KM effector cell; protease receptor-1; EPR-1; embryonic tissue; tubulin;  
 KM mitotic spindle; abnormal cell growth; tumorigenesis; ss.

XX Homo sapiens.

OS WO950440-A2.

PN 07-OCT-1999.

PD 01-APR-1999; 99WO-US007205.

PF 01-APR-1998; 98US-0080288P.

PR (UYVA ) UNITV YALE.

PA Altieri DC;

PI WPI; 1999-S91329/50.

DR Identifying agents that modulate Survivin interactions.

PS Example; Fig 3A; 56pp; English.

XX This sequence is the 5' flanking region of the human Survivin gene,  
 CC (Survivin open reading frame: AAZ31037). Survivin is a 142 amino acid  
 CC protein of approximately 16.5kD. Survivin is a member of the IAP family  
 CC of apoptosis inhibitors, and the gene is located on chromosome 17q25. The  
 CC nucleic acid sequence of Survivin is related to the Effector Cell  
 CC Protease Receptor-1 (EPR-1). Survivin is expressed in tumour cells and  
 CC embryonic tissue. The sequence is used in the invention which involves  
 CC identifying an agent that modulates interactions between Survivin and  
 CC tubulin. Agents that modulate interaction between Survivin and tubulin or  
 CC mitotic spindles may be administered to a mammal to modulate biological  
 CC or pathological processes mediated by Survivin, for example Survivin-  
 CC mediated inhibition of cellular apoptosis. Thus an inhibiting agent may  
 CC be used to block abnormal cell growth, for example during tumorigenesis.  
 CC An agent that increases interaction may be used to extend cell growth in  
 CC culture

XX Sequence 403 BP; 63 A; 143 C; 140 G; 57 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.28e-30 Length: 403  
 Score: 37.00 Matches: 37  
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 26.06% Indels: 0  
DB: 2 Gaps: 0  
US-09-690-825-34 (1-142) x AA231036 (1-403)  
QY 1 MetGlyAlaProThreLeuProProAlaTyrGlnProPheLeuLysAspHisArgTleSer 20  
DB 291 ATGGGTGCCCCGACGTTGGCCCTGCTGCAGCCCTTTCTCAAGACCCGCACTCTCT 350  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArg 37  
DB 351 ACATTCAAGAACTGGCCCTTTCTTGAGGGCTGCGCTGCACCCCGGAGCGG 401  
RESULT 29  
AA82458  
ID AA82458 standard; DNA; 1100 BP.  
XX  
AC AA82458;  
XX  
XX 29-UTN-2001 (first entry)  
DT  
XX Human survivin promoter.  
XX  
XX Human; Map4 promoter; cytosratic; gene therapy; apoptosis modulator;  
KM survivin promoter; p53-mediated repression; cancer;  
KM proliferative disorder; ds.  
XX  
XX Homo sapiens.  
OS  
XX MO200123522-A2.  
PN  
XX 05-APR-2001.  
PD  
XX 02-OCT-2000; 2000MO-US027078.  
PF  
XX 30-SEP-1999; 99US-0157171P.  
PR  
XX (FOX- ) FOX CHASE CANCER CENT.  
PA  
XX  
PI Murphy M;  
XX  
DR WPI; 2001-308140/32.  
XX  
XX Novel nucleic acid sequence comprising a promoter region containing p53  
PT negative response elements are useful for treating cancers and other  
PT cellular proliferative diseases.  
XX  
XX Claim 15; Page 68-69; 84pp; English.  
PS  
XX The present sequence is the human survivin promoter. The invention  
CC relates to isolated nucleic acid molecules containing the Map4 or  
CC survivin promoter. The Map4 and survivin genes are involved in p53-  
CC mediated repression of gene expression. The sequences are useful for  
CC determining the presence of an alteration in the Map 4 or survivin  
CC promoter sequence, where an alteration is associated with a pathological  
CC condition. They are also useful for screening novel therapeutic agents  
CC for treating and diagnosing cancer and other cellular proliferative  
CC disorders. Furthermore, the sequences may be used in gene therapy  
XX  
SQ Sequence 1100 BP; 199 A; 343 C; 312 G; 246 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1.58e-28 Length: 1100  
Score: 36.00 Matches: 36  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 25.35% Indels: 0  
DB: 4 Gaps: 0  
US-09-690-825-34 (1-142) x AA82458 (1-1100)  
QY 1 MetGlyAlaProThreLeuProProAlaTyrGlnProPheLeuLysAspHisArgTleSer 20

DB 991 ATGGGTGCCCCGACGTTGGCCCTGCTGCAGCCCTTTCTCAAGACCCGCACTCTCT 1050  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGlu 36  
DB 1051 ACATTCAAGAACTGGCCCTTTCTTGAGGGCTGCGCTGCACCCCGGAG 1058  
RESULT 30  
ADC32487/c  
ID ADC32487 standard; cDNA; 1596 BP.  
XX  
AC ADC32487;  
XX  
DT 18-DEC-2003 (first entry)  
DT  
XX Human novel cDNA contig sequence, SEQ ID NO:2569.  
XX  
XX Human; diagnostic; drug screening; forensic; gene mapping;  
KM biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
KM neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
KM ulcers; osteoporosis; autoimmune disease; cancer;  
KM molecular weight marker; food supplement; antiparkinsonian; nootropic;  
KM neuroprotective; antihaemic; anticoagulant; thrombolytic; vulnerary;  
KM anticancer; osteopathic; immunosuppressive; antiinflammatory; cycostatic;  
KM gene therapy; chromosome 17; ss.  
XX  
XX Homo sapiens.  
OS  
XX MO2003029271-A2.  
PN  
XX 10-APR-2003.  
PD  
XX 24-SEP-2002; 2002MO-US030474.  
PF  
XX 24-SEP-2001; 2001US-0324631P.  
PR  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;  
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Wang G;  
PI Haley-Vicence D, Dymnae RT;  
XX  
XX WPI; 2003-371981/35.  
DR P-PSDB; ADC33254.  
XX  
XX New polynucleotide and polypeptide useful for diagnosing, preventing or  
PT treating conditions such as neurodegenerative diseases, anaemias, platelet  
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
PT cancer.  
XX  
XX Example 2; SEQ ID NO 2569; 1185pp; English.  
PS  
XX The invention relates to 971 novel human cDNA sequences (ADC29919-  
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
CC invention also relates to nucleic acid sequences over 99% identical with  
CC the novel human cDNAs. The invention additionally encompasses expression  
CC vectors and host cells comprising a nucleic acid of the invention; the  
CC recombinant production of a polypeptide of the invention; an antibody  
CC against a polypeptide of the invention; a method of detecting  
CC polynucleotides or polypeptides of the invention; and methods of  
CC identifying a compound which binds to a polypeptide of the invention. The  
CC invention further discloses methods of preventing, treating or  
CC ameliorating a medical condition; kits comprising polynucleotide probes  
CC and/or monoclonal antibodies for carrying out the methods of the  
CC invention; methods for the identification of compounds that modulate the  
CC expression or activity of the polynucleotide and/or polypeptide; and 767  
CC contig sequences corresponding to the cDNA sequences of the invention  
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
CC -ADC33394). The nucleic acids and polypeptides of the invention are  
CC useful in diagnostics, drug screening, forensic, gene mapping, in the  
CC identification of mutations responsible for genetic disorders or other  
CC traits, for assessing biodiversity, and in producing many other types of  
CC data and products dependent on DNA and amino acid sequences. They are

CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
CC disease and other neurodegenerative diseases, anemia, platelet  
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
CC cancer. The nucleic acids may also be used as hybridisation probes or  
CC primers, and in the recombinant production of a protein. The polypeptides  
CC are also useful in generating antibodies, as molecular weight markers,  
CC and as food supplements. The present sequence represents a human contig  
CC sequence used in an example of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1596 BP, 328 A, 509 C, 470 G, 289 T, 0 U, 0 Other;

Alignment Scores:

Pred. No.:	4,87e-06	Length:	1596
Score:	15.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.56%	Indels:	0
DB:	9	Gaps:	0

US-09-690-825-34 (1-142) x ADC32487 (1-1596)

Qy 114 AAlaySGlUTtRrARrAsnLyLyGClUphGtUgUThrAla 128  
|||  
Db 90 GCAAGAGAACCAACATTAAGAGAGAAATTGAGGAAACTGCG 46

RESULT 31  
ABK93852  
ID ABK93852 standard; DNA; 30 BP.

XX ABK93852;

DT 26-AUG-2002 (first entry)

DE Human survivin Real Time-PCR probe.

XX Human, sb; antisense; inhibitor of apoptosis; H1A1, H1A2, XIAP;  
KM cytoskeletal; cancer; ovarian cancer; adenocarcinoma; lymphoma; IAP;  
KM pancreatic cancer; embryonic development; viral pathogenesis;  
KM autoimmune disorder; neurodegenerative disease; multiple sclerosis;  
KM lupus erythematosus; herpes virus infection; pox virus infection;  
KM adenovirus infection; proliferative disease; probe; real time PCR.

XX Homo sapiens.

OS WO200226968-A2.

PN 04-APR-2002.

PD 27-SEP-2001; 2001WO-CA001379.

PF 28-SEP-2000; 2000US-00672717.

PR (UYOT-) UNIV OTTAWA.

PA (ABGE-) ABERGIA THERAPEUTICS INC.

XX Korneluk RG, Lacasse E, Baird S, Holcik M, Young S;  
PI WPI; 2002-479562/51.

DR Novel antisense inhibitor of apoptosis nucleic acid useful for enhancing  
PT apoptosis in a cell, for treating cancer and other proliferative  
PT diseases.

XX Example 4; Page 42; 135pp; English.

XX The invention relates to an inhibitor of apoptosis (IAP) antisense  
CC nucleic acid (I) that inhibits IAP biological activity, regardless of  
CC length of the antisense nucleic acid, the IAP proteins may be mouse or  
CC human XIAP, H1A1 or H1A2. Also included are a pharmaceutical  
CC composition comprising a mammalian IAP antisense molecule and a method of

CC enhancing apoptosis in a cell, comprising administering a negative  
CC regulator of the IAP anti-apoptotic pathway to the cell. The IAP  
CC antisense inhibitor is useful for enhancing apoptosis in a cell in a  
CC mammal diagnosed with a proliferative disease. The method is useful for  
CC treating a patient diagnosed with a proliferative disease like cancer.

CC The IAP antisense molecule is useful to treat, ameliorate, improve,  
CC sustain or prevent proliferative diseases (e.g. ovarian cancer,  
CC adenocarcinoma, lymphoma, pancreatic cancer,) and also in diseases or  
CC conditions where apoptosis is involved or implicated (e.g. embryonic  
CC development, viral pathogenesis, autoimmune disorders, neurodegenerative  
CC diseases, multiple sclerosis, lupus erythematosus and infection by herpes  
CC virus, pox virus and adenovirus). The present sequence is a real time PCR  
CC probe used to measure mRNA levels in an experiment showing that the  
CC antisense molecules of the invention reduce the levels of IAP mRNA in a  
CC cell

XX SQ Sequence 30 BP, 12 A, 8 C, 7 G, 3 T, 0 U, 0 Other;

Alignment Scores:

Pred. No.:	0.276	Length:	30
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.34%	Indels:	0
DB:	6	Gaps:	0

US-09-690-825-34 (1-142) x ABK93852 (1-30)

Qy 69 ProASPASPpProlleGluGluHis 77  
|||  
Db 3 CCAGATGACGACCCCATAGAGAGACAT 29

RESULT 32  
AAS21533  
ID AAS21533 standard; DNA; 26 BP.

XX AAS21533;

DT 21-NOV-2001 (first entry)

DE Mouse survivin PCR probe.

XX Survivin; human; mouse; cytoskeletal; antisense oligonucleotide;  
KM hyperproliferative condition; cancer; apoptosis; cytokinesis; ss.

XX Mus musculus.

OS WO200157059-A1.

PN 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US002939.

PF 02-FEB-2000; 2000US-00496694.

PR (ISIS-) ISIS PHARM INC.

XX Bennett CF, Ackermann EJ, Swayze EE, Cowsett LM;  
PI WPI; 2001-48863/53.

DR Novel antisense compounds for modulating the expression of Survivin and  
PT treatment of cancer.

XX Example 13; Page 50; 120pp; English.

XX The invention relates to antisense oligonucleotides targeted to a nucleic  
CC acid molecule encoding human Survivin, where the antisense  
CC oligonucleotide inhibits the expression of human Survivin. These  
CC antisense oligonucleotides are used in the treatment of an animal  
CC suffering from a disease or condition associated with Survivin, e.g. a  
CC hyperproliferative condition such as cancer, and comprises administering  
CC a therapeutically or prophylactically effective amount of the antisense

CC oligonucleotide so that expression of Survivin is inhibited. The  
CC oligonucleotides can also be used to treat a human suffering from a  
CC disease or condition characterised by a reduction in apoptosis comprising  
CC administering the antisense oligonucleotide to a human. In addition, the  
CC antisense oligonucleotide and a cytotoxic chemotherapeutic agent e.g.  
CC taxol or cisplatin, can be used to modulate apoptosis, cytokinesis or the  
CC cell cycle, or inhibit the proliferation in a cancer cell by contacting  
CC the cell with the antisense oligonucleotide. AAS21521-AAS21768 represent  
CC Survivin nucleic acids, and antisense oligonucleotides targeted to  
CC Survivin, used in the method of the invention

XX  
SQ Sequence 26 BP; 8 A; 4 C; 10 G; 4 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.79 Length: 26  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: Gaps: 0

US-09-690-825-34 (1-142) x AAS21533 (1-26)

QY 62 LysGluLeuGluGlyTyrGluPro 69  
ABV52787  
ID AAV52787 standard; cDNA; 509 BP.  
AC AAV52787;  
XX  
XX 17-SEP-2002 (first entry)  
DT  
XX  
DE Human prostate expression marker cDNA 52778.  
XX  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX pharmacogenomic marker; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX MO200160860-A2.  
XX  
XX 23-AUG-2001.  
XX  
XX 20-FEB-2001; 2001WO-US005171.  
XX  
XX 17-FEB-2000; 2000US-0183319P.  
XX 16-MAR-2000; 2000US-0189862P.  
XX 25-MAY-2000; 2000US-0207454P.  
XX 09-JUN-2000; 2000US-0211314P.  
XX 18-JUL-2000; 2000US-0219007P.  
XX 13-DEC-2000; 2000US-0255281P.  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX Schlegel R, Endege WO, Monahan JB;  
XX WPI, 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
XX prostate cells and correlating with presence of prostate cancer; useful  
XX for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 10233; 11750pp; English.

XX  
XX The invention relates to an isolated nucleic acid molecule (1) comprising  
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
XX specification or its complement. (1) is useful for: (a) assessing whether  
XX a patient is afflicted with prostate cancer; (b) monitoring the efficacy  
XX progression of prostate cancer in a patient; (c) assessing the efficacy  
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or incidence of prostate cancer in a patient  
CC; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX  
SQ Sequence 509 BP; 212 A; 110 C; 73 G; 114 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 46.3 Length: 509  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: Gaps: 0

US-09-690-825-34 (1-142) x ABV52787 (1-509)

QY 115 LysGluThrAsnAsnLysLysLys 122  
ABV52787  
ID AAV52787 standard; cDNA; 515 BP.  
AC AAV52787;  
XX  
XX 21-NOV-2001 (first entry)  
DT  
XX  
DE DNA encoding mouse survivin.  
XX  
XX Survivin; human; mouse; cytostatic; antisense oligonucleotide;  
XX hyperproliferative condition; cancer; apoptosis; cytokinesis; ss.  
XX  
XX Mus musculus.  
XX  
XX MO200157059-A1.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US002939.  
XX  
XX 02-FEB-2000; 2000US-00496694.  
XX  
XX (ISIS-) ISIS PHARM INC.  
XX  
XX Bennett CF, Ackermann EJ, Swayze EE, Cowser LM;  
XX WPI, 2001-48863/53.  
XX  
XX Novel antisense compounds for modulating the expression of Survivin and  
XX treatment of cancer.

PS Example 18; Page 116-117; 120pp; English.

XX  
XX The invention relates to antisense oligonucleotides targeted to a nucleic  
XX acid molecule encoding human Survivin, where the antisense  
XX oligonucleotide inhibits the expression of human Survivin. These  
XX antisense oligonucleotides are used in the treatment of an animal  
XX suffering from a disease or condition associated with Survivin, e.g. a  
XX hyperproliferative condition such as cancer, and comprises administering  
XX a therapeutically or prophylactically effective amount of the antisense  
XX oligonucleotide so that expression of Survivin is inhibited. The  
XX oligonucleotides can also be used to treat a human suffering from a  
XX disease or condition characterised by a reduction in apoptosis comprising  
XX administering the antisense oligonucleotide to a human. In addition, the  
XX antisense oligonucleotide and a cytotoxic chemotherapeutic agent e.g.  
XX taxol or cisplatin, can be used to modulate apoptosis, cytokinesis or the  
XX cell cycle, or inhibit the proliferation in a cancer cell by contacting  
XX the cell with the antisense oligonucleotide. AAS21521-AAS21768 represent  
XX Survivin nucleic acids, and antisense oligonucleotides targeted to

CC Survivin, used in the method of the invention  
XX Sequence 515 BP, 152 A, 120 C, 109 G, 133 T, 0 U, 1 Other;  
SQ Sequence 515 BP, 152 A, 120 C, 109 G, 133 T, 0 U, 1 Other;  
Alignment Scores:  
Pred. No.: 46.8 Length: 515  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: Gaps: 0  
US-09-690-825-34 (1-142) x AAS21751 (1-515)  
QY 122 LysGluPheGluGluThrAlaLys 129  
DB 325 AAAGGTTTGAAGAGACTGCAAG 348  
RESULT 35  
ACH27866  
ID ACH27866 standard; cDNA; 587 BP.  
XX ACH27866;  
AC ACH27866;  
XX 13-OCT-2003 (first entry)  
DT 13-OCT-2003 (first entry)  
XX Human adult ovary cDNA #6246.  
DE Human adult ovary cDNA #6246.  
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
KM genome mapping; biodiversity; genetic disorder.  
XX Homo sapiens.  
OS Homo sapiens.  
XX US2003073623-A1.  
PN 17-APR-2003.  
PD 17-APR-2003.  
XX 30-JUN-2001; 2001US-00918995.  
PF 30-JUN-2001; 2001US-00918995.  
XX 30-JUN-2001; 2001US-00918995.  
PR 30-JUN-2001; 2001US-00918995.  
XX (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
XX (STRC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
XX (JONE/) JONES L W.  
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX WPI; 2003-615964/58.  
DR WPI; 2003-615964/58.  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
PS Claim 1; SEQ ID NO 15078; 44bp; English.  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from USPRO at  
CC seqdata.uspro.gov/sequence.html?docid=20030073623  
XX Sequence 587 BP, 133 A, 150 C, 110 G, 91 T, 0 U, 103 Other;  
SQ Sequence 587 BP, 133 A, 150 C, 110 G, 91 T, 0 U, 103 Other;  
Alignment Scores:  
Pred. No.: 52.9 Length: 587  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: Gaps: 0  
US-09-690-825-34 (1-142) x ACH27866 (1-587)  
QY 1 MetGlyAlaProThrLeuProPro 8  
DB 367 ATGGAGCCCCCACACTCCCTCCT 390  
RESULT 36  
ACA01764  
ID ACA01764 standard; DNA; 660 BP.  
XX ACA01764;  
AC ACA01764;  
XX 03-JUN-2003 (first entry)  
DT 03-JUN-2003 (first entry)  
XX C. glutamicum derived ORF SEQ ID 1755.  
DE C. glutamicum derived ORF SEQ ID 1755.  
XX Coryneform; nucleic acid array; fermentation; culture; ds.  
KM Coryneform; nucleic acid array; fermentation; culture; ds.  
XX Corynebacterium glutamicum.  
OS Corynebacterium glutamicum.  
XX DE10128510-A1.  
PN 19-DEC-2002.  
PD 19-DEC-2002.  
XX 13-JUN-2001; 2001DE-01028510.  
PF 13-JUN-2001; 2001DE-01028510.  
XX 13-JUN-2001; 2001DE-01028510.  
PR 13-JUN-2001; 2001DE-01028510.  
XX (DEGS) DEGUSSA AG.  
PA Farwick M, Moeckel B, Pfeifferle W, Bathe B, Hutmacher K;  
XX WPI; 2003-279970/28.  
DR WPI; 2003-279970/28.  
XX New nucleic acid array useful for monitoring mRNA expression of  
PT Corynebacterium glutamicum during fermentation, comprising nucleic acid  
PT from Corynebacterium glutamicum.  
XX Claim 1; Page 590; 709pp; German.  
PS This invention describes a novel nucleic acid array involving  
CC Corynebacterium glutamicum polynucleotides. The arrays are used to  
CC analyse C. glutamicum, particularly for monitoring a fermentation process  
CC to determine expression levels of C. glutamicum cellular mRNA. Such  
CC monitoring particularly differentiates between expression levels of  
CC different strains of C. glutamicum and allows the adjustment of different  
CC culture and fermentation conditions. ACA00010-ACA02188 represent C.  
CC glutamicum derived polynucleotides described in the disclosure of the  
CC invention.  
XX Sequence 660 BP, 183 A, 190 C, 141 G, 146 T, 0 U, 0 Other;  
SQ Sequence 660 BP, 183 A, 190 C, 141 G, 146 T, 0 U, 0 Other;  
Alignment Scores:  
Pred. No.: 59.1 Length: 660  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: Gaps: 0

US-09-690-825-34 (1-142) x ACA01764 (1-660)

QY 4 ProthleupProProAlatrgln 11  
 |||||  
 DB 152 CCGACCTTCCCCCTGCATGCAG 175

RESULT 37  
 AAC69515  
 ID AAC69515 standard; DNA; 1441 BP.  
 XX  
 AC AAC69515;  
 XX  
 DT 31-JAN-2001 (first entry)  
 XX  
 DE Human secreted protein gene 4 clone HKJUK47.  
 XX  
 KM Immunosuppressive; antiarthritic; antirheumatic; antiproliferative;  
 KM cytostatic; cardiatic; vasotropic; cerebroprotective; neuroprotective;  
 KM nootropic; antibacterial; virucide; fungicide; ophtalmological;  
 KM vulnerary; gene therapy; infection; human; secreted protein; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200061623-A1.  
 XX  
 PD 19-OCT-2000.  
 XX  
 PF 06-APR-2000; 2000MO-US008979.  
 XX  
 PR 09-APR-1999; 99US-0128693P.  
 PR 26-APR-1999; 99US-0130991P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Ni J, Komatsculis CA, Rosen CA, Sceppe DR, Shi Y,  
 PI Lafleur DW, Olsen HS, Ebner R, Florence KA, Moore PA, Birse CB,  
 PI Young PE;  
 XX  
 DR WPI; 2000-647418/62.  
 XX  
 PT New nucleic acid molecules encoding 62 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions and  
 PT used as food additives or preservatives.  
 XX  
 PS Claim 1; Page 531; 716pp; English.  
 XX  
 CC The invention relates to the isolation of genes (AAC69512-C69587)  
 CC encoding 62 human secreted proteins (AAB38321-B38396). The genes can be  
 CC used to generate fusion proteins by linking to the gene for the human  
 CC immunoglobulin G Fc portion (AAC69503) for increasing the stability of  
 CC the fusion protein as compared to the human protein only. The genes and  
 CC proteins are useful for preventing, ameliorating or treating medical  
 CC conditions, e.g. by protein or gene therapy. The genes are isolated from  
 CC a range of human tissues disclosed in the specification. The nucleic  
 CC acids, proteins, antibodies and (ant)agonists are useful in the  
 CC diagnosis, treatment and prevention of: (a) autoimmune diseases e.g.  
 CC rheumatoid arthritis; (b) hyperproliferative disorders e.g. neoplasms of  
 CC the breast or liver; (c) cardiovascular disorders e.g. cardiac arrest;  
 CC (d) cerebrovascular disorders e.g. cerebral ischemia; (e) angioneurosis;  
 CC (f) nervous system disorders e.g. Alzheimer's disease; (g) infections  
 CC caused by bacteria, viruses and fungi; and (h) ocular disorders e.g.  
 CC corneal infection. The polypeptides can also be used to aid wound healing  
 CC and epithelial cell proliferation, to prevent skin aging due to sunburn,  
 CC to maintain organs before transplantation, for supporting cell culture of  
 CC primary tissues, to regenerate tissues and in chemotaxis  
 XX  
 SQ Sequence 1441 BP; 348 A; 367 C; 503 G; 221 T; 0 U; 2 Other;  
 XX

Alignment Scores:  
 Pred. No.: 123 Length: 1441  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 5.63% Indels: 0  
 DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x AAC69515 (1-1441)

QY 1 MetGlyAlaProthleupProPro 8  
 |||||  
 DB 1343 ATGGAGACCCCACTCTCTCT 1366

RESULT 38  
 ABS97149/c  
 ID ABS97149 standard; DNA; 1469 BP.  
 XX  
 AC ABS97149;  
 XX  
 DT 23-DEC-2002 (first entry)  
 XX  
 DE Wild type human CYP450A1 gene.  
 XX  
 KM Human; ds; cytochrome P450 A1; CYP4501A1; UGT2B4; MDR1;  
 KM cytochrome P450 A2; CYP4501A2; cytochrome P450 02E; CYP45002E1; LTF;  
 KM adrenergic receptor beta1; ADRB1; aryl hydrocarbon; AHR; MRE3; NR112;  
 KM aryl hydrocarbon receptor nuclear translocator; ARNT; cathepsin S; CTSS;  
 KM cyclooxygenase 2; COX2; diazepam binding inhibitor; DBI; haematological;  
 KM epoxide hydroxylase 2; EPHX2; 5-lipoxygenase activating protein; FLAP;  
 KM glutathione-S-transferase 12; GST12; histamine-N-methyl transferase;  
 KM HMMT; kallikrein 2; KLK2; nicotinamide-N-methyl transferase; NMMT;  
 KM NADPH quinone oxidoreductase 2; NQO2; sulfoxyltransferase thermolabile; STM;  
 KM UDP-glucuronosyl transferase 2B4; UDP-glucuronosyl transferase 2B7;  
 KM UGT2B7; UDP-glucuronosyl transferase; UGT2B15; urokinase receptor; uPA;  
 KM multidrug resistance 1; lactotransferrin; orphan nuclear receptor;  
 KM acetylcholine muscarinic receptor; CHMR1; CHMR2; CHMR3; CHMR4; CHMR5;  
 KM altered drug metabolism; cardiovascular function; colorectal tumour;  
 KM central nervous system; pulmonary; immunological; SNP;  
 KM single nucleotide polymorphism.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200257410-A2.  
 XX  
 PD 25-JUL-2002.  
 XX  
 PF 28-NOV-2001; 2001MO-US044838.  
 XX  
 PR 28-NOV-2000; 2000US-00724389.  
 XX  
 PA (DNAS-) DNA SCI LAB INC.  
 XX  
 PI Guida M, Hall J;  
 XX  
 DR WPI; 2002-698522/75.  
 XX  
 PT Isolated nucleic acid molecules having polymorphisms in known human genes  
 PT e.g. cytochrome p450 and cathepsin S useful as genetic linkage markers  
 PT for locating, identifying and characterizing the genes responsible for  
 PT disorder-related traits.  
 XX  
 PS Example 1; Page 209; 714pp; English.  
 XX  
 CC This invention relates to the sequence of an isolated nucleic acid  
 CC molecule comprising at least one base variation from that of a known  
 CC human cytochrome P450 A1 (CYP4501A1), cytochrome P450 A2 (CYP4501A2),  
 CC cytochrome P450 02E1 (CYP45002E1), adrenergic receptor beta1 (ADRB1),  
 CC aryl hydrocarbon (AHR), aryl hydrocarbon receptor nuclear translocator  
 CC (ARNT), cathepsin S (CTSS), cyclooxygenase 2 (COX2), diazepam binding  
 CC inhibitor (DBI), epoxide hydroxylase 2 (EPHX2), 5-lipoxygenase activating  
 CC protein (FLAP), glutathione-S-transferase 12 (GST12), histamine-N-methyl  
 CC transferase (HMMT), kallikrein 2 (KLK2), nicotinamide-N-methyl  
 CC transferase (NMMT), NADPH quinone oxidoreductase 2 (NQO2),  
 CC sulfoxyltransferase thermolabile (STM), UDP-glucuronosyl transferase 2B4  
 CC (UGT2B4), UDP-glucuronosyl transferase 2B7 (UGT2B7), UDP-glucuronosyl  
 CC transferase (UGT2B15), urokinase receptor (uPA), multidrug resistance 1

CC (MDR1), lactotransferrin (LTF), multidrug resistance associated protein 3  
 CC (MRP3), orphan nuclear receptor (NR112), or acetylcholine muscarinic  
 CC receptor 1, 2, 3, 4, or 5 (CHMR1, CHMR2, CHMR3, CHMR4 or CHMR5) sequence.  
 CC The polymorphisms in the human genes cited in the invention are useful as  
 CC genetic linkage markers for locating and characterizing the genes that  
 CC are responsible for specific traits within the genome and eventually  
 CC identifying the genes responsible for a variety of disorder-related  
 CC traits as a result of their e.g., overexpression, constitutive  
 CC expression, mutation or underexpression, which may be used in diagnosing  
 CC and/or treating the disorders. The nucleic acid molecules comprising the  
 CC polymorphic sequences contained in CYP4501A1, CYP4501A2, CYP4501A3, AHR,  
 CC ARNT, EPH2, GST12, NNMT, NQO2, NR12, STM, UGT2B4, UGT2B15, AHR,  
 CC MDR1 and/or MDR3 are useful for screening individuals for altered drug  
 CC metabolism. The polymorphic sequences contained in CYP4501A1, CYP4501A2,  
 CC AHR, MDR1 and/or MDR3 may also be used to screen individuals for  
 CC susceptibility to cancer. Polymorphic sequences in ADRB1 or CHMR2 are  
 CC used to screen for altered cardiovascular function, in COX2 for altered  
 CC susceptibility to colorectal tumors, in DBI or CHMR1 for altered central  
 CC nervous system function, in FLAP and HMT for altered pulmonary,  
 CC immunological or haematological function, in KIK2 for altered serine  
 CC protease activity in the prostate, in LTF for altered immunological or  
 CC haematological function, in CHMR3, CHMR4 or CHMR5 for altered central and  
 CC peripheral nervous system function. The present sequence represents a  
 CC polymorphic DNA sequence of the invention

XX SQ Sequence 1469 BP; 245 A; 497 C; 457 G; 270 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	126	Length:	1469
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	6	Gaps:	0

US-09-690-825-34 (1-142) x ABS97149 (1-1469)

Oy 2 G1YAlaProThreProProAla 9

DB 820 GAGGCCCCACCCCTACCCCGGCT 797

RESULT 39

AAZ43802 ID AAZ43802 standard; cDNA; 1493 BP.

XX AAZ43802;

DT 10-MAR-2000 (first entry)

XX Human adult skin cDNA clone vcl\_1.

XX Human; secreted protein; treatment; nutritional activity; cytokine;

XX cell proliferation; cell differentiation; hematopoiesis regulation;

XX tissue growth; activin; inhibitor; chemotactic; chemokine; hemostatic;

XX thrombolytic; anti-inflammatory; invasion suppressor; tumor inhibition;

XX gene therapy; ds.

XX Homo sapiens.

XX WO955721-A1.

XX 04-NOV-1999.

XX 23-APR-1999; 99WO-US008504.

XX 24-APR-1998; 98US-0082904P.

XX 11-JUN-1998; 98US-0088994P.

XX 12-JUN-1998; 98US-0089278P.

XX 02-JUL-1998; 98US-0091647P.

XX 24-AUG-1998; 98US-0097639P.

XX 22-APR-1999; 99US-00097639.

XX (ALPH-) ALPHAGENE INC.

XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapleyko P;  
 PI :  
 XX WPI; 2000-052801/04.  
 DR P-PSDB; AAY50939, AAY50940.

XX New polynucleotides encoding secreted human proteins, derived from human  
 PT fetal brain, adult skin, adult brain, adult heart, adult thymus and adult  
 PT aorta cDNA libraries.

XX Claim 60a; Page 252-253; 282pp; English.

XX This invention describes novel human secreted proteins which are encoded  
 CC by polynucleotides obtained from fetal brain, adult skin, adult brain,  
 CC adult heart, adult thymus and adult aorta cDNA libraries. The  
 CC polynucleotides and proteins are predicted to have biological activities  
 CC which would make them suitable for treating, preventing or ameliorating  
 CC medical conditions in humans and animals, although no supporting data is  
 CC given. Suggested activities include nutritional activity, cytokine and  
 CC cell proliferation/differentiation activity, immune stimulating (e.g. as  
 CC vaccines) or suppressing activity, hematopoiesis regulating activity,  
 CC tissue growth activity, activin/inhibin activity,  
 CC chemotactic/chemokinetic activity, hemostatic and thrombolytic activity,  
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumor  
 CC invasion suppressor activity, and tumor inhibition activity. The  
 CC polynucleotides are also stated to be useful for gene therapy. AAZ43777-  
 CC Z43808 represent the polynucleotides described in the invention which  
 CC encode the proteins represented in AAY50905-Y50947

XX SQ Sequence 1493 BP; 429 A; 356 C; 492 G; 216 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	128	Length:	1493
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	3	Gaps:	0

US-09-690-825-34 (1-142) x AAZ43802 (1-1493)

Oy 1 MetG1YAlaProThreProPro 8

DB 1312 ATGGAGCCCCCAGACTCCCTCT 1335

RESULT 40

AAAZ7036 ID AAAZ7036 standard; cDNA; 1733 BP.

XX AAAZ7036;

DT 08-AUG-2000 (first entry)

XX Human PRO1411 (UNQ729) cDNA sequence SEQ ID NO:51.

XX Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;

XX transmembrane; secretion; immunoadhesion; pharmaceutical; screening; ss.

XX Homo sapiens.

XX WO200012708-A2.

XX 09-MAR-2000.

XX 01-SEP-1999; 99WO-US020111.

XX 01-SEP-1998; 98US-0098716P.

XX 01-SEP-1998; 98US-0098749P.

XX 01-SEP-1998; 98US-0098750P.

XX 02-SEP-1998; 98US-0098803P.

XX 02-SEP-1998; 98US-0098821P.

XX 02-SEP-1998; 98US-0098843P.

XX 09-SEP-1998; 98US-0099536P.

PR	09-SEP-1998;	98US-0099536P.
PR	09-SEP-1998;	98US-0099538P.
PR	09-SEP-1998;	98US-0099602P.
PR	09-SEP-1998;	98US-0099642P.
PR	10-SEP-1998;	98US-0099741P.
PR	10-SEP-1998;	98US-0099754P.
PR	10-SEP-1998;	98US-0099763P.
PR	10-SEP-1998;	98US-0099792P.
PR	10-SEP-1998;	98US-0099808P.
PR	10-SEP-1998;	98US-0099812P.
PR	10-SEP-1998;	98US-0099815P.
PR	10-SEP-1998;	98US-0099816P.
PR	15-SEP-1998;	98US-0100385P.
PR	15-SEP-1998;	98US-0100388P.
PR	15-SEP-1998;	98US-0100390P.
PR	16-SEP-1998;	98US-0100584P.
PR	16-SEP-1998;	98US-0100627P.
PR	16-SEP-1998;	98US-0100661P.
PR	16-SEP-1998;	98US-0100662P.
PR	16-SEP-1998;	98US-0100664P.
PR	17-SEP-1998;	98US-0100683P.
PR	17-SEP-1998;	98US-0100700P.
PR	17-SEP-1998;	98US-0100710P.
PR	17-SEP-1998;	98US-0100711P.
PR	17-SEP-1998;	98US-0100919P.
PR	17-SEP-1998;	98US-0100930P.
PR	18-SEP-1998;	98US-0100848P.
PR	18-SEP-1998;	98US-0100849P.
PR	18-SEP-1998;	98US-0101014P.
PR	18-SEP-1998;	98US-0101068P.
PR	18-SEP-1998;	98US-0101071P.
PR	22-SEP-1998;	98US-0101279P.
PR	22-SEP-1998;	98US-0101471P.
PR	23-SEP-1998;	98US-0101472P.
PR	23-SEP-1998;	98US-0101474P.
PR	23-SEP-1998;	98US-0101475P.
PR	23-SEP-1998;	98US-0101476P.
PR	23-SEP-1998;	98US-0101477P.
PR	23-SEP-1998;	98US-0101479P.
PR	24-SEP-1998;	98US-0101738P.
PR	24-SEP-1998;	98US-0101741P.
PR	24-SEP-1998;	98US-0101743P.
PR	24-SEP-1998;	98US-0101915P.
PR	24-SEP-1998;	98US-0101916P.
PR	24-SEP-1998;	98US-0102207P.
PR	29-SEP-1998;	98US-0102240P.
PR	29-SEP-1998;	98US-0102307P.
PR	29-SEP-1998;	98US-0102330P.
PR	29-SEP-1998;	98US-0102484P.
PR	30-SEP-1998;	98US-0102487P.
PR	30-SEP-1998;	98US-0102570P.
PR	30-SEP-1998;	98US-0102571P.
PR	01-OCT-1998;	98US-0102684P.
PR	01-OCT-1998;	98US-0102687P.
PR	02-OCT-1998;	98US-0102965P.
PR	06-OCT-1998;	98US-0103258P.
PR	06-OCT-1998;	98US-0103449P.
PR	07-OCT-1998;	98US-0103314P.
PR	07-OCT-1998;	98US-0103315P.
PR	07-OCT-1998;	98US-0103328P.
PR	07-OCT-1998;	98US-0103395P.
PR	07-OCT-1998;	98US-0103396P.
PR	07-OCT-1998;	98US-0103401P.
PR	08-OCT-1998;	98US-0103633P.
PR	08-OCT-1998;	98US-0103678P.
PR	08-OCT-1998;	98US-0103719P.
PR	14-OCT-1998;	98US-0104257P.
PR	20-OCT-1998;	98US-0104987P.
PR	20-OCT-1998;	98US-0105000P.
PR	20-OCT-1998;	98US-0105002P.
PR	21-OCT-1998;	98US-0105104P.
PR	22-OCT-1998;	98US-0105169P.
PR	22-OCT-1998;	98US-0105266P.
PR	26-OCT-1998;	98US-0105639P.
PR	26-OCT-1998;	98US-0105694P.
PR	27-OCT-1998;	98US-0105807P.
PR	27-OCT-1998;	98US-0105881P.
PR	27-OCT-1998;	98US-0105882P.
PR	27-OCT-1998;	98US-0106062P.
PR	28-OCT-1998;	98US-0106023P.
PR	28-OCT-1998;	98US-0106029P.
PR	28-OCT-1998;	98US-0106030P.
PR	28-OCT-1998;	98US-0106032P.
PR	28-OCT-1998;	98US-0106033P.
PR	28-OCT-1998;	98US-0106178P.
PR	29-OCT-1998;	98US-0106248P.
PR	29-OCT-1998;	98US-0106384P.
PR	30-OCT-1998;	98US-0106500P.
PR	03-NOV-1998;	98US-0106856P.
PR	03-NOV-1998;	98US-0106902P.
PR	03-NOV-1998;	98US-0106905P.
PR	03-NOV-1998;	98US-0106912P.
PR	03-NOV-1998;	98US-0106933P.
PR	10-NOV-1998;	98US-0107833P.
PR	17-NOV-1998;	98US-0108775P.
PR	17-NOV-1998;	98US-0108779P.
PR	17-NOV-1998;	98US-0108787P.
PR		



```

Query Match:      5.63%      Indels:      0
DB:               3          Gaps:         0
US-09-690-825-34 (1-142) x AAF54238 (1-1733)
Cy               1 MetGlyAlaProThreProPro 8
Db              1534 ATGGAGCCCCCAGCTCCTCCT 1557

RESULT 41
AAF54238
ID AAF54238 standard; DNA; 1733 BP.
AC AAF54238;
XX
XX 02-APR-2001 (first entry)
DT
DE DNA encoding protein of the invention #15.
XX
XX Secreted; transmembrane; gene therapy; ss.
XX
XX Unidentified.
XX
XX WO200078961-A1.
XX
XX 28-DEC-2000.
XX
XX 18-FEB-2000; 2000WO-US004342.
XX
XX 23-JUN-1999; 99US-0141037P.
XX 20-JUL-1999; 99US-0144758P.
XX 26-JUL-1999; 99US-0145698P.
XX 01-SEP-1999; 99WO-US020111.
XX 29-OCT-1999; 99US-0162506P.
XX 30-NOV-1999; 99WO-US028313.
XX 02-DEC-1999; 99WO-US028551.
XX 16-DEC-1999; 99WO-US030095.
XX 05-JAN-2000; 2000WO-US000219.
XX 06-JAN-2000; 2000WO-US000376.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
XX Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
XX Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
XX Williams PM, Wood WI;
XX WPI; 2001-071395/08.
XX
XX Secreted and transmembrane proteins and nucleic acids designated PRO,
XX useful as hybridization probes, in chromosome and gene mapping and gene
XX therapy.
XX
XX Claim 2; Fig 29; 787pp; English.
XX
XX The present invention relates to secreted and transmembrane proteins.
XX These proteins and the DNA encoding them may be used as hybridization
XX probes, in chromosome and gene mapping and in the generation of anti-
XX sense RNA and DNA. They may also be used to generate either
XX transgenic animals or knockout animals which are in turn useful for
XX development and screening of therapeutically useful reagents. The nucleic
XX acids may also be used in gene therapy
XX
XX SQ Sequence 1733 BP; 502 A; 397 C; 580 G; 254 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.:      147
Score:         8.00      Length:      1733
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     5.63%      Indels: 0
DB:              4          Gaps:      0

```

```

US-09-690-825-34 (1-142) x AAF54238 (1-1733)
Cy               1 MetGlyAlaProThreProPro 8
Db              1534 ATGGAGCCCCCAGCTCCTCCT 1557

RESULT 42
AAC87038
ID AAC87038 standard; cDNA; 1734 BP.
AC AAC87038;
XX
XX 20-APR-2001 (first entry)
DT
DE Nucleotide sequence of human polypeptide PRO1411.
XX
XX Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;
XX PRO185; PRO210; PRO215; PRO217; PRO242; PRO288; PRO365; PRO1361; PRO1308;
XX PRO1183; PRO1272; PRO1419; PRO4959; PRO7170; PRO248; PRO353; PRO1318;
XX PRO1600; PRO9940; PRO533; PRO301; PRO187; PRO337; PRO1411; PRO4356;
XX PRO246; PRO265; PRO941; PRO10096; PRO6003; PRO6004; PRO350; PRO2630;
XX PRO6309; cell death; genetic disorder; transgenic animal; gene therapy;
XX ss.
XX
XX Homo sapiens.
XX
XX Key      Location/Qualifiers
XX CDS      184..1506
XX          /*tag= a
XX          184..246
XX          /*tag= b
XX
XX WO200077037-A2.
XX
XX 21-DEC-2000.
XX
XX 22-MAY-2000; 2000WO-US014042.
XX
XX 15-JUN-1999; 99US-0139695P.
XX 20-JUL-1999; 99US-0145070P.
XX 26-JUL-1999; 99US-0145698P.
XX 17-AUG-1999; 99US-0149396P.
XX 01-SEP-1999; 99WO-US020111.
XX 08-SEP-1999; 99WO-US020594.
XX 15-SEP-1999; 99WO-US021090.
XX 15-SEP-1999; 99WO-US021547.
XX 30-NOV-1999; 99WO-US028313.
XX 01-DEC-1999; 99WO-US028301.
XX 02-DEC-1999; 99WO-US028565.
XX 07-DEC-1999; 99US-0169495P.
XX 05-JAN-2000; 2000WO-US000219.
XX 18-FEB-2000; 2000WO-US004341.
XX 18-FEB-2000; 2000WO-US004342.
XX 22-FEB-2000; 2000WO-US004414.
XX 01-MAR-2000; 2000WO-US005601.
XX 02-MAR-2000; 2000WO-US005841.
XX 20-MAR-2000; 2000WO-US007377.
XX 30-MAR-2000; 2000WO-US008439.
XX 15-MAY-2000; 2000WO-US013358.
XX 17-MAY-2000; 2000WO-US013705.
XX
XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein DA, Desnoyers L, Eaton DL;
XX Ferrara N, Fong S, Gao W, Garber H, Gerritsen ME, Goddard A;
XX Godowski PJ, Gurney AL, Kijavini JV, Mather JP, Napier MA, Pan J;
XX Paoni NF, Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM;
XX Wood WI, Zhang Z;
XX WPI; 2001-050091/06.
XX P-FSDB; AAB31205.
XX
XX Isolated nucleic acid molecule encoding a PRO polypeptide which is a

```

PT transmembrane polypeptide is useful for gene therapy and identification of related polypeptides.

PS Claim 2; Fig 53; 24pp; English.

XX The present sequence encodes a human secreted and transmembrane  
CC polypeptide. The specification describes human polypeptides, designated  
CC PRO186, PRO444, PRO183, PRO185, PRO210, PRO215, PRO217, PRO242, PRO288,  
CC PRO365, PRO1361, PRO1308, PRO1183, PRO1272, PRO1419, PRO4999, PRO7170,  
CC PRO248, PRO353, PRO1318, PRO1600, PRO5940, PRO513, PRO301, PRO187,  
CC PRO337, PRO1411, PRO4356, PRO246, PRO265, PRO941, PRO10996, PRO6003,  
CC PRO6004, PRO350, PRO2630 and PRO6309. The biological activity of cells  
CC can be modulated with agents that bind to these polypeptides, resulting  
CC in the death of the cells. The polynucleotides encoding these  
CC polypeptides are useful in the recombinant production of the  
CC polypeptides, as a hybridisation probe to screen libraries to isolate  
CC homologous sequences, or to map the gene. They may also be used for  
CC analysing genetic disorders, and to produce transgenic animals which are  
CC useful for the development and screening of therapeutically useful  
CC reagents. The polynucleotides can also be used in gene therapy e.g. to  
CC replace a defective gene

XX Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	147	Length:	1734
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	4	Gaps:	0

US-09-690-825-34 (1-142) x AAC87038 (1-1734)

OY 1 MetGlyAlaProthrleuProPro 8

DB 1534 ATGGAGACCCCCACACTCCTCTCT 1557

#### RESULT 43

ID AAS46025 standard; cDNA; 1734 BP.

XX AAS46025;

AC 18-DEC-2001 (first entry)

DE Human DNA encoding PRO polypeptide sequence #101.

XX PRO polypeptide: mammal; tumour; cancer; human; cattle; horse; sheep; ss;

KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;

KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;

KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;

XX PCR primer.

XX Homo sapiens.

PN WO200168848-A2.

PD 20-SEP-2001.

XX 28-FEB-2001; 2001WO-US006520.

XX 01-MAR-2000; 2000WO-US005601.

PR 02-MAR-2000; 2000WO-US005841.

PR 03-MAR-2000; 2000US-0187202P.

PR 06-MAR-2000; 2000US-0186588P.

PR 14-MAR-2000; 2000US-0189320P.

PR 14-MAR-2000; 2000US-0189328P.

PR 15-MAR-2000; 2000WO-US006884.

PR 21-MAR-2000; 2000US-0190828P.

PR 21-MAR-2000; 2000US-0191007P.

PR 21-MAR-2000; 2000US-0191048P.

PR 21-MAR-2000; 2000US-0191314P.

PR 28-MAR-2000; 2000US-0192655P.  
PR 29-MAR-2000; 2000US-0193032P.  
PR 29-MAR-2000; 2000US-0193053P.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 04-APR-2000; 2000US-0194449P.  
PR 04-APR-2000; 2000US-0194647P.  
PR 11-APR-2000; 2000US-0195975P.  
PR 11-APR-2000; 2000US-0196000P.  
PR 11-APR-2000; 2000US-0196187P.  
PR 11-APR-2000; 2000US-0196590P.  
PR 11-APR-2000; 2000US-0196820P.  
PR 18-APR-2000; 2000US-0198121P.  
PR 18-APR-2000; 2000US-0198585P.  
PR 25-APR-2000; 2000US-0199397P.  
PR 25-APR-2000; 2000US-0199550P.  
PR 25-APR-2000; 2000US-0199654P.  
PR 03-MAY-2000; 2000US-0201516P.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 05-JUN-2000; 2000US-0209832P.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 22-AUG-2000; 2000US-00644848.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000WO-US034956.

XX (GENTH ) GENENTECH INC.

XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;

XX Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

XX WPI; 2001-602746/68.

XX P-PSDB; AAU29124.

XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the

XX presence of tumors, such as prostate and breast tumors, in mammals and to

XX screen for modulators of the compounds.

XX Claim 2; Fig 201; 774pp; English.

XX Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR  
CC primers for PRO polypeptides of the invention. The sequences of the  
CC invention can be used to detect the presence of a tumour in a mammal by  
CC comparing the level of expression of a PRO polypeptide in a test sample  
CC of cells from the animal and a control sample of normal cells, whereby a  
CC higher level of expression in the test sample indicates the presence of a  
CC tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,  
CC pigs, goats and rabbits but are preferably human. The polypeptides can be  
CC used to stimulate tumour necrosis factor (TNF) alpha release from human  
CC blood. When contacted with it, A specific polypeptide can be used to  
CC stimulate the proliferation or differentiation of chondrocyte cells. The  
CC PRO proteins can be used to determine the presence of tumours and also  
CC susceptibility to tumour development, particularly adrenal, lung, colon,  
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian  
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids  
CC can be used for genetic analysis of individuals with genetic disorders

XX Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	147	Length:	1734
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	4	Gaps:	0

US-09-690-825-34 (1-142) x AAS46025 (1-1734)

OY 1 MetGlyAlaProthrleuProPro 8

Db 1534 ATGGAGCCCCCACACTCCCTCCT 1557

RESULT 44  
ID AAF92083 standard; cDNA; 1734 BP.

AAF92083;

15-MAY-2001 (first entry)

Human PRO1411 cDNA.

Human; PRO protein; mapping; ss.

Homo sapiens.

WO200116318-A2.

08-MAR-2001.

24-AUG-2000; 2000WO-US023328.

01-SEP-1999; 99WO-US020111.

13-SEP-1999; 99WO-US021030.

07-DEC-1999; 99US-0169495P.

09-DEC-1999; 99US-0170262P.

11-JAN-2000; 2000US-0175481P.

18-FEB-2000; 2000WO-US004341.

18-FEB-2000; 2000WO-US004342.

22-FEB-2000; 2000WO-US004414.

01-MAR-2000; 2000WO-US005601.

21-MAR-2000; 2000US-0187022P.

30-MAR-2000; 2000WO-US008439.

25-APR-2000; 2000US-0199397P.

22-MAY-2000; 2000WO-US014042.

05-JUN-2000; 2000US-0209832P.

(GETH ) GENENTECH INC.

Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;

Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;

WPI; 2001-183260/18.

P-PSDB; AAB87551.

Eighty four nucleic acids encoding PRO polypeptides, useful in molecular

biology, including use as hybridization probes, and in chromosome and

gene mapping.

Claim 2; Fig 51; 278pp; English.

Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;

Alignment Scores: 147

Pred. No.: 8.00

Score: 100.00%

Percent Similarity: 100.00%

Best Local Similarity: 5.63%

Query Match: 4

DB: 0

US-09-690-825-34 (1-142) x AAF92083 (1-1734)

1 MetGlyAlaProThrLeuProPro 8

ABK33613 standard; cDNA; 1734 BP.

ABK33613;

08-MAY-2002 (first entry)

CDNA encoding human PRO protein, Seq ID No 155.

Human; secreted protein; PRO; tumour; lung cancer; colon cancer;

breast cancer; prostate tumour; rectal tumour; liver tumour;

pericyte cell proliferation; chondrocyte cell proliferation;

tumour necrosis factor-alpha; gene; ss.

Homo sapiens.

WO200208288-A2.

31-JAN-2002.

29-JUN-2001; 2001WO-US021066.

20-JUL-2000; 2000US-0219556P.

25-JUL-2000; 2000US-0220585P.

25-JUL-2000; 2000US-0220605P.

25-JUL-2000; 2000US-0220607P.

25-JUL-2000; 2000US-0220624P.

25-JUL-2000; 2000US-0220638P.

25-JUL-2000; 2000US-0220664P.

25-JUL-2000; 2000US-0220666P.

26-JUL-2000; 2000US-0220839P.

28-JUL-2000; 2000WO-US020710.

01-AUG-2000; 2000US-0222425P.

22-AUG-2000; 2000US-0227133P.

23-AUG-2000; 2000WO-US023522.

24-AUG-2000; 2000WO-US023328.

10-NOV-2000; 2000WO-US030873.

28-NOV-2000; 2000US-0253646P.

01-DEC-2000; 2000WO-US032678.

20-DEC-2000; 2000US-00747259.

20-DEC-2000; 2000WO-US034956.

28-FEB-2001; 2001WO-US006520.

01-MAR-2001; 2001WO-US006666.

22-MAR-2001; 2001US-00816744.

10-MAY-2001; 2001US-00854208.

10-MAY-2001; 2001US-00854280.

25-MAY-2001; 2001WO-US017092.

(GETH ) GENENTECH INC.

Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

WPI; 2002-172001/22.

P-PSDB; AAB83669.

One hundred and twenty two nucleic acids encoding PRO polypeptides,

useful for treating a PRO related disorder and for diagnosing tumors such

as lung cancer, colon cancer, breast tumor, prostate tumor, rectal tumor

or liver tumor.

Claim 2; Fig 155; 359pp; English.

The invention relates to one hundred and twenty two nucleic acids

encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides

encode human secreted proteins. The PRO nucleic acids, polypeptides,

CC agonists and antagonists are useful for treating a PRO related disorder.  
CC The PRO polypeptides are useful for diagnosing tumours, especially lung  
CC cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or  
CC liver tumour. The PRO polypeptides are useful for stimulating the  
CC proliferation of, or gene expression, in pericyte cells, for stimulating  
CC the proliferation or differentiation of chondrocyte cells, for  
CC stimulating the release of tumour necrosis factor-alpha from human blood,  
CC for stimulating or inhibiting the proliferation of normal human dermal  
CC fibroblast cells. The PRO polypeptide may also be used as molecular  
CC weight markers and for tissue typing. The PRO nucleic acids have  
CC applications in molecular biology, including use as hybridisation probes,  
CC and in chromosome and gene mapping. ABK3536-ABK3657 represent human PRO  
CC protein coding sequences of the invention

SQ Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 147 Length: 1734  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: Gaps: 0

US-09-690-825-34 (1-142) x ABK33613 (1-1734)

QY 1 MetGlyAlaProThreupPro 8  
Db 1534 ATGGAGCCCCCAGCTCCCTCCT 1557

RESULT 46  
ABK74403  
ID ABK74403 standard; CDNA; 1734 BP.  
AC ABK74403;  
XX  
XX  
DT 10-DEC-2002 (first entry)  
XX  
XX  
DE Human cDNA encoding secreted/transmembrane protein PRO1411.  
XX  
XX  
KW Human; ss; gene; secreted protein; transmembrane protein; antirheumatic;  
KW antiarthritic; osteopathic; sports-related joint problem;  
KW articular cartilage defect; osteoarthritis; rheumatoid arthritis.  
XX  
XX  
OS Homo sapiens.  
XX  
XX  
PN US2002119130-A1.  
XX  
PD 29-AUG-2002.  
XX  
PF 06-DEC-2001; 2001US-00006667.  
XX  
XX  
PR 29-OCT-1997; 97US-006335P.  
PR 29-OCT-1997; 97US-0064215P.  
PR 22-APR-1998; 98US-0082797P.  
PR 29-APR-1998; 98US-0083495P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 04-JUN-1998; 98US-0088621P.  
PR 04-JUN-1998; 98US-0088629P.  
PR 04-JUN-1998; 98US-0088630P.  
PR 10-JUN-1998; 98US-0088734P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090346P.  
PR 24-JUN-1998; 98US-0090444P.

PR 25-JUN-1998; 98US-0090668P.  
PR 25-JUN-1998; 98US-0090669P.  
PR 26-JUN-1998; 98US-0090862P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 18-AUG-1998; 98US-0096549P.  
PR 18-AUG-1998; 98US-0096599P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097979P.  
PR 01-SEP-1998; 98US-0098749P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099792P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 10-SEP-1998; 98US-0099815P.  
PR 16-SEP-1998; 98US-0100627P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0101330P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 22-SEP-1998; 98US-0101279P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101916P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 08-MAR-1999; 99MO-US005028.  
PR 14-MAY-1999; 99MO-US010733.  
PR 02-JUN-1999; 99MO-US012252.  
PR 01-SEP-1999; 99MO-US020111.  
PR 15-SEP-1999; 99MO-US021090.  
PR 15-SEP-1999; 99MO-US021194.  
PR 22-DEC-1999; 99MO-US030720.  
PR 18-FEB-2000; 2000MO-US004341.  
PR 18-FEB-2000; 2000MO-US004342.  
PR 22-FEB-2000; 2000MO-US004414.  
PR 01-MAR-2000; 2000MO-US005601.  
PR 30-MAR-2000; 2000MO-US008439.  
PR 22-MAY-2000; 2000MO-US014042.  
PR 02-JUN-2000; 2000MO-US015264.  
PR 23-AUG-2000; 2000MO-US023522.  
PR 24-AUG-2000; 2000MO-US023528.  
PR 10-NOV-2000; 2000MO-US030873.  
PR 01-DEC-2000; 2000MO-US032378.  
PR 20-DEC-2000; 2000MO-US034956.  
PR 28-FEB-2001; 2001MO-US006520.  
PR 01-MAR-2001; 2001MO-US006666.  
PR 30-MAY-2001; 2001MO-US017443.  
PR 01-JUN-2001; 2001MO-US017800.  
PR 20-JUN-2001; 2001MO-US019692.  
PR 29-JUN-2001; 2001MO-US021066.  
PR 09-JUL-2001; 2001MO-US021735.  
XX  
XX  
PA (GETH ) GENENTECH INC.  
XX  
XX  
PI Baton DL, Filvarcoff E, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;  
XX  
XX  
DR WPI: 2002-731348/79.  
DR P-PSDB; ABG95876.  
XX  
XX  
PT New isolated secreted and transmembrane PRO polypeptide useful for  
PT modulating biological activity of a cell, or for treating sports-related  
PT joint problems, osteoarthritis or rheumatoid arthritis.  
XX  
XX  
PS Claim 2; Fig 51; 399pp; English.  
XX  
XX  
CC The invention relates to an isolated secreted and transmembrane PRO  
CC polypeptide having 80 % sequence identity to a sequence appearing as

ABG95851-ABG95934 or their associated signal peptide, or a sequence of an extracellular domain of the proteins with their associated signal peptide or lacking its associated signal peptide. Also included are the nucleic acids encoding the proteins, vectors, host cells, fusion proteins and antibodies which specifically bind to the proteins. The proteins are useful for detecting a polypeptide designated as A, B, C or D in a sample suspected of containing an A, B, C or D polypeptide, by contacting the sample with a polypeptide designated as E, F, G, H or I (or vice versa) and determining the formation of A/E, B/F, B/G, C/H or D/I polypeptide conjugate in the sample, where the formation of the conjugate is indicative of the presence of an A, B, C or D polypeptide in the sample, where A is a PRO10272 polypeptide, B is a PRO210 polypeptide, C is a PRO10096 polypeptide, D is a PRO19760 polypeptide, E is a PRO05601 polypeptide, F is a PRO1 polypeptide, G is a PRO20040 polypeptide, H is a PRO20233 polypeptide and I is a PRO1890 polypeptide. The sample comprises a cell suspected of expressing the A, B, C or D polypeptide. The E, F, G, H or I polypeptide is labeled with a detectable label or is attached to a solid support. The proteins are useful for linking a bioactive molecule to a cell expressing a polypeptide designated as A, B, C or D or E, F, G, H or I. The bioactive molecule is a toxin, a radiolabel or an antibody. The bioactive molecule causes death of the cell. A, B, C, D, E, F, G, H, or I, or antibodies against them are useful for modulating a biological activity of a cell expressing a polypeptide designated as A, B, C or D or E, F, G, H, or I. The cell is killed. The proteins are useful for identifying agonists or antagonists, for the preparation of a medicament useful in the treatment of a condition which is responsive to the proteins, as molecular weight markers for protein electrophoresis purposes, and as therapeutic agents for treating sports-related joint problems, articular cartilage defects, osteoarthritis or rheumatoid arthritis. Nucleic acids encoding the proteins are useful as hybridisation probes, in chromosome and gene mapping, in the generation of anti-sense RNA and DNA, for the preparation of the proteins, to generate transgenic or knockout animals which are useful in the development and screening of therapeutic useful reagents, for chromosome identification, and in gene therapy. The antibody is useful as a therapeutic agent, in a diagnostic assay and for affinity purification of the protein from recombinant cell culture natural sources. The present sequence encodes a novel secreted or transmembrane protein of the invention.

SEQ Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	147	Length:	1734
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	6	Gaps:	0

US-09-690-825-34 (1-142) X ABG74403 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8

DB 1534 ATGGAGGCCCACTCCCTCCT 1557

RESULT 47  
ABL95635  
ID ABL95635 standard; cDNA; 1734 BP.

AC ABL95635;

DT 19-JUL-2002 (first entry)

DE Human angiogenesis related cDNA PRO1411 SEQ ID NO: 149.

XX Human; angiogenesis; PRO protein; cardiovascularisation; wound; cancer;  
XX atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder;  
KM cardiatic; cytostatic; antiangiogenic; hypotensive; vulnerary;  
KM antiarteriosclerotic; gene; ss.  
XX  
OS Homo sapiens.  
XX

PN WO200208284-A2.  
XX 31-JAN-2002.  
XX 09-JUL-2001; 2001WO-US021735.  
XX 20-JUL-2000; 2000US-0219556P.  
XX 25-JUL-2000; 2000US-0220624P.  
XX 25-JUL-2000; 2000US-022064P.  
XX 28-JUL-2000; 2000WO-US020710.  
XX 02-AUG-2000; 2000US-0222695P.  
XX 17-AUG-2000; 2000US-00643657.  
XX 23-AUG-2000; 2000WO-US023522.  
XX 24-AUG-2000; 2000WO-US023328.  
XX 07-SEP-2000; 2000US-0230978P.  
XX 18-SEP-2000; 2000US-0064610.  
XX 18-SEP-2000; 2000US-0066350.  
XX 24-OCT-2000; 2000US-0242922P.  
XX 08-NOV-2000; 2000US-00709238.  
XX 10-NOV-2000; 2000WO-US030952.  
XX 10-NOV-2000; 2000WO-US030952.  
XX 01-DEC-2000; 2000WO-US032678.  
XX 20-DEC-2000; 2000US-00747259.  
XX 20-DEC-2000; 2000WO-US034956.  
XX 22-JAN-2001; 2001US-00767609.  
XX 28-FEB-2001; 2001US-00796498.  
XX 28-FEB-2001; 2001WO-US006520.  
XX 01-MAR-2001; 2001WO-US006666.  
XX 09-MAR-2001; 2001US-00802706.  
XX 14-MAR-2001; 2001US-0080689.  
XX 22-MAR-2001; 2001US-00816744.  
XX 05-APR-2001; 2001US-00828366.  
XX 10-MAY-2001; 2001US-00854208.  
XX 25-MAY-2001; 2001US-00866029.  
XX 25-MAY-2001; 2001US-00866034.  
XX 25-MAY-2001; 2001WO-US017092.  
XX 30-MAY-2001; 2001US-00870574.  
XX 30-MAY-2001; 2001WO-US017443.  
XX 01-JUN-2001; 2001WO-US017800.  
XX 20-JUN-2001; 2001WO-US019692.  
XX (GERTH ) GENENTECH INC.  
XX (BAKE) BAKER K P.  
XX (FERR) FERRARA N.  
XX (GERB) GERBER H.  
XX (GERR) GERRTSEN M E.  
XX (GODD) GODDARD A.  
XX (GODD) GODDARD P J.  
XX (GURN) GURNEY A L.  
XX (HILL) HILLAN K J.  
XX (MARS) MARSTERS S A.  
XX (PANT) PAN J.  
XX (PAON) PAONI N F.  
XX (SREP) STEPHAN J F.  
XX (WATA) WATANABE C K.  
XX (WILL) WILLIAMS P M.  
XX (WOOD) WOOD W I.  
XX Baker KP, Ferrara N, Gerber H, Gerltzen ME, Goddard A,  
XX Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF,  
XX Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W,  
XX WPI; 2002-171999/22.  
XX P-PSDB; ABB95497.  
XX

One hundred and eighty seven nucleic acids encoding PRO polypeptides, useful in diagnosis and treatment of cardiovascular (e.g. myocardial infarction), endothelial or angiogenic disorders in a mammal.

Claim 1; Fig 149; 567pp; English.

The present invention provides the protein and coding sequences of human

CC PRO proteins. These are useful for treating or diagnosing a  
CC cardiovascular, endometrial or angiogenic disorder, including cardiac  
CC hypertrophy, trauma, cancer, age-related macular degeneration,  
CC atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,  
CC angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour  
CC angiogenesis (such as breast carcinoma and liver carcinoma) and wound  
CC healing. The present sequence is a coding sequence of the invention  
XX

Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	147	Length:	1734
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	6	Gaps:	0

US-09-690-825-34 (1-142) x ABL95635 (1-1734)

Oy 1 MetGlyAlaProThrLeuProPro 8

Db 1534 ATGGAGCCCCCAGCTCCTCCT 1557

RESULT 48

ABX78628 standard; cDNA; 1734 BP.

XX ABX78628;

AC ABX78628;

XX

DT 15-APR-2003 (first entry)

XX

DE Human PRO polynucleotide #101.

XX

XX Human: PRO; gene; ss; cytosolic; tumour; cancer; breast; lung; stomach;  
KM liver; dog; cat; cow; horse; sheep; pig; goat; rabbit; ADBPT;  
KM antibody-dependent enzyme mediated prodng therapy.

XX

OS Homo sapiens.

XX

PN US2003027272-A1.

XX

PD 06-FEB-2003.

XX

PF 21-JUN-2002; 2002US-00176492.

XX

PR 18-SEP-1997; 97US-0059263P.

PR 18-SEP-1997; 97US-0059266P.

PR 17-OCT-1997; 97US-0062250P.

PR 21-OCT-1997; 97US-0063486P.

PR 24-OCT-1997; 97US-0063120P.

PR 24-OCT-1997; 97US-0063121P.

PR 28-OCT-1997; 97US-0063540P.

PR 28-OCT-1997; 97US-0063541P.

PR 28-OCT-1997; 97US-0063544P.

PR 28-OCT-1997; 97US-0063564P.

PR 29-OCT-1997; 97US-0063734P.

PR 31-OCT-1997; 97US-0063870P.

PR 31-OCT-1997; 97US-0064103P.

PR 13-NOV-1997; 97US-0065111P.

PR 21-NOV-1997; 97US-0066120P.

PR 24-NOV-1997; 97US-0066466P.

PR 24-NOV-1997; 97US-0066722P.

PR 11-DEC-1997; 97US-0069335P.

PR 12-DEC-1997; 97US-0069425P.

PR 17-DEC-1997; 97US-0069670P.

PR 18-DEC-1997; 97US-0068017P.

PR 10-MAR-1998; 98US-0077450P.

PR 11-MAR-1998; 98US-0077632P.

PR 11-MAR-1998; 98US-0077649P.

PR 20-MAR-1998; 98US-0078865P.

PR 20-MAR-1998; 98US-0078839P.

PR 27-MAR-1998; 98US-0079664P.

PR 27-MAR-1998; 98US-0079786P.

PR 31-MAR-1998; 98US-0080107P.

PR 01-APR-1998; 98US-0080194P.

PR 01-APR-1998; 98US-0080337P.

PR 01-APR-1998; 98US-0080333P.

PR 08-APR-1998; 98US-0081049P.

PR 08-APR-1998; 98US-0081070P.

PR 09-APR-1998; 98US-0081195P.

PR 15-APR-1998; 98US-0081838P.

PR 21-APR-1998; 98US-0082568P.

PR 22-APR-1998; 98US-0082569P.

PR 22-APR-1998; 98US-0082704P.

PR 28-APR-1998; 98US-0082797P.

PR 28-APR-1998; 98US-0083372P.

PR 29-APR-1998; 98US-0083455P.

PR 29-APR-1998; 98US-0083496P.

PR 29-APR-1998; 98US-0083499P.

PR 29-APR-1998; 98US-0083559P.

PR 29-APR-1998; 98US-0083592P.

PR 05-MAY-1998; 98US-0084366P.

PR 06-MAY-1998; 98US-0084414P.

PR 07-MAY-1998; 98US-0084639P.

PR 07-MAY-1998; 98US-0084640P.

PR 07-MAY-1998; 98US-0084643P.

PR 15-MAY-1998; 98US-0085579P.

PR 15-MAY-1998; 98US-0085580P.

PR 15-MAY-1998; 98US-0085582P.

PR 15-MAY-1998; 98US-0085700P.

PR 18-MAY-1998; 98US-0086023P.

PR 22-MAY-1998; 98US-0086392P.

PR 22-MAY-1998; 98US-0086486P.

PR 28-MAY-1998; 98US-0087098P.

PR 28-MAY-1998; 98US-0087208P.

PR 02-JUN-1998; 98US-0087609P.

PR 02-JUN-1998; 98US-0087759P.

PR 03-JUN-1998; 98US-0087827P.

PR 04-JUN-1998; 98US-0088025P.

PR 04-JUN-1998; 98US-0088028P.

PR 04-JUN-1998; 98US-0088029P.

PR 04-JUN-1998; 98US-0088033P.

PR 04-JUN-1998; 98US-0088326P.

PR 05-JUN-1998; 98US-0088167P.

PR 05-JUN-1998; 98US-0088202P.

PR 05-JUN-1998; 98US-0088212P.

PR 05-JUN-1998; 98US-0088217P.

PR 09-JUN-1998; 98US-0088655P.

PR 10-JUN-1998; 98US-0088722P.

PR 10-JUN-1998; 98US-0088738P.

PR 10-JUN-1998; 98US-0088811P.

PR 10-JUN-1998; 98US-0088824P.

PR 10-JUN-1998; 98US-0088825P.

PR 10-JUN-1998; 98US-0088826P.

PR 11-JUN-1998; 98US-0088861P.

PR 11-JUN-1998; 98US-0088863P.

PR 11-JUN-1998; 98US-0088876P.

PR 12-JUN-1998; 98US-0089090P.

PR 12-JUN-1998; 98US-0089105P.

PR 16-JUN-1998; 98US-0089512P.

PR 16-JUN-1998; 98US-0089514P.

PR 17-JUN-1998; 98US-0089518P.

PR 17-JUN-1998; 98US-0089598P.

PR 17-JUN-1998; 98US-0089653P.

PR 17-JUN-1998; 98US-0089658P.

PR 18-JUN-1998; 98US-0089908P.

PR 19-JUN-1998; 98US-0089952P.

PR 22-JUN-1998; 98US-0090246P.

PR 22-JUN-1998; 98US-0090252P.

PR 22-JUN-1998; 98US-0090254P.

PR 24-JUN-1998; 98US-0090429P.

PR 24-JUN-1998; 98US-0090435P.

PR 24-JUN-1998; 98US-0090444P.

PR 24-JUN-1998; 98US-0090461P.

PR 24-JUN-1998; 98US-0090535P.

PR 24-JUN-1998; 98US-0090540P.

```
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095988P.
PR 10-AUG-1998; 98US-0096012P.
PR 10-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096955P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100652P.
PR 16-SEP-1998; 98US-0100654P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101930P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.

PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Alignment Scores:
Pred. No.: 147
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.63%
DB: 7
Gaps: 0

US-09-690-825-34 (1-142) x ABX78628 (1-1734)
QY 1 MetGlyValaProThrLeuProPro 8
Db 1534 ATGGAGCCCCCACACTCCCTCTCT 1557

RESULT 49
ACA75600
ID ACA75600 standard; cDNA; 1734 BP.
XX
AC ACA75600;
XX
DT 07-JUL-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1411 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; gene therapy;
KW tumour necrosis factor-alpha release; TNF-alpha release;
KW chondrocyte proliferation; chondrocyte differentiation; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003032127-A1.
XX
PD 13-FEB-2003.
XX
PF 26-JUN-2002; 2002US-00183012.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063421P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 29-OCT-1997; 97US-0063564P.
PR 31-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
```





PR 30-SEP-1998; 98US-0102571P.  
 PR 01-OCT-1998; 98US-0102684P.  
 PR 01-OCT-1998; 98US-0102687P.  
 PR 02-OCT-1998; 98US-0102965P.  
 PR 06-OCT-1998; 98US-0103258P.

## Alignment Scores:

Pred. No.: 147 Length: 1734  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.63% Indels: 0  
 DB: 7 Gaps: 0

US-09-690-825-34 (1-142) x ACA75600 (1-1734)

Qy 1 MetGlyAlaProThrLeuPro 8

D5 1534 ATGGAGCCGCCACACCTCTCT 1557

## RESULT 50

ACA71080

ID ACA71080 standard; cDNA; 1734 BP.

XX AC ACA71080;

XX DT 02-AUG-2003 (first entry)

XX DE Human secreted/transmembrane protein (PRO) cDNA #101.

XX KW Human; gene; ss; secreted and transmembrane protein; PRO; TNF-alpha;  
 tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;  
 tissue typing.

XX OS Homo sapiens.

XX FN US2003032112-A1.

XX PD 13-FEB-2003.

XX PF 21-JUN-2002; 2002US-00176756.

PR 18-SEP-1997; 97US-0059263P.

PR 18-SEP-1997; 97US-0059266P.

PR 17-OCT-1997; 97US-0062250P.

PR 21-OCT-1997; 97US-0063486P.

PR 24-OCT-1997; 97US-0063120P.

PR 24-OCT-1997; 97US-0063121P.

PR 28-OCT-1997; 97US-0063540P.

PR 28-OCT-1997; 97US-0063541P.

PR 28-OCT-1997; 97US-0063544P.

PR 29-OCT-1997; 97US-0063734P.

PR 31-OCT-1997; 97US-0063870P.

PR 31-OCT-1997; 97US-0064103P.

PR 13-NOV-1997; 97US-0065311P.

PR 21-NOV-1997; 97US-0066120P.

PR 24-NOV-1997; 97US-0066466P.

PR 11-DEC-1997; 97US-0066772P.

PR 12-DEC-1997; 97US-0069435P.

PR 17-DEC-1997; 97US-0069870P.

PR 18-DEC-1997; 97US-0068017P.

PR 10-MAR-1998; 98US-0077450P.

PR 11-MAR-1998; 98US-0077632P.

PR 11-MAR-1998; 98US-0077649P.

PR 20-MAR-1998; 98US-0078886P.

PR 20-MAR-1998; 98US-0078939P.

PR 27-MAR-1998; 98US-0079664P.

PR 27-MAR-1998; 98US-0079786P.

PR 31-MAR-1998; 98US-0080187P.

PR 31-MAR-1998; 98US-0080194P.

PR 01-APR-1998; 98US-0080327P.

PR 01-APR-1998; 98US-0080333P.  
 PR 08-APR-1998; 98US-0081049P.  
 PR 08-APR-1998; 98US-0081070P.  
 PR 09-APR-1998; 98US-0081195P.  
 PR 15-APR-1998; 98US-0081838P.  
 PR 21-APR-1998; 98US-0082568P.  
 PR 21-APR-1998; 98US-0082569P.  
 PR 22-APR-1998; 98US-0082704P.  
 PR 22-APR-1998; 98US-0082797P.  
 PR 28-APR-1998; 98US-0083322P.  
 PR 29-APR-1998; 98US-0083495P.  
 PR 29-APR-1998; 98US-0083496P.  
 PR 29-APR-1998; 98US-0083499P.  
 PR 29-APR-1998; 98US-0083559P.  
 PR 05-MAY-1998; 98US-0084366P.  
 PR 07-MAY-1998; 98US-0084414P.  
 PR 07-MAY-1998; 98US-0084639P.  
 PR 07-MAY-1998; 98US-0084840P.  
 PR 15-MAY-1998; 98US-0085579P.  
 PR 15-MAY-1998; 98US-0085580P.  
 PR 15-MAY-1998; 98US-0085582P.  
 PR 15-MAY-1998; 98US-0085700P.  
 PR 18-MAY-1998; 98US-0086023P.  
 PR 22-MAY-1998; 98US-0086392P.  
 PR 22-MAY-1998; 98US-0086486P.  
 PR 28-MAY-1998; 98US-0087098P.  
 PR 02-JUN-1998; 98US-0087609P.  
 PR 03-JUN-1998; 98US-0087759P.  
 PR 04-JUN-1998; 98US-0087827P.  
 PR 04-JUN-1998; 98US-0088028P.  
 PR 04-JUN-1998; 98US-0088029P.  
 PR 04-JUN-1998; 98US-0088033P.  
 PR 04-JUN-1998; 98US-0088126P.  
 PR 05-JUN-1998; 98US-0088167P.  
 PR 05-JUN-1998; 98US-0088202P.  
 PR 05-JUN-1998; 98US-0088212P.  
 PR 05-JUN-1998; 98US-0088217P.  
 PR 09-JUN-1998; 98US-0088555P.  
 PR 10-JUN-1998; 98US-0088722P.  
 PR 10-JUN-1998; 98US-0088738P.  
 PR 10-JUN-1998; 98US-0088740P.  
 PR 10-JUN-1998; 98US-0088811P.  
 PR 10-JUN-1998; 98US-0088824P.  
 PR 10-JUN-1998; 98US-0088825P.  
 PR 10-JUN-1998; 98US-0088826P.  
 PR 11-JUN-1998; 98US-0088861P.  
 PR 11-JUN-1998; 98US-0088863P.  
 PR 11-JUN-1998; 98US-0088876P.  
 PR 12-JUN-1998; 98US-0089090P.  
 PR 12-JUN-1998; 98US-0089105P.  
 PR 16-JUN-1998; 98US-0089512P.  
 PR 17-JUN-1998; 98US-0089514P.  
 PR 17-JUN-1998; 98US-0089538P.  
 PR 17-JUN-1998; 98US-0089598P.  
 PR 17-JUN-1998; 98US-0089653P.  
 PR 18-JUN-1998; 98US-0089908P.  
 PR 19-JUN-1998; 98US-0089952P.  
 PR 22-JUN-1998; 98US-0090246P.  
 PR 22-JUN-1998; 98US-0090252P.  
 PR 22-JUN-1998; 98US-0090254P.  
 PR 24-JUN-1998; 98US-0090429P.  
 PR 24-JUN-1998; 98US-0090435P.  
 PR 24-JUN-1998; 98US-0090444P.  
 PR 24-JUN-1998; 98US-0090461P.  
 PR 24-JUN-1998; 98US-0090535P.  
 PR 24-JUN-1998; 98US-0090540P.  
 PR 25-JUN-1998; 98US-0090576P.  
 PR 25-JUN-1998; 98US-0090578P.  
 PR 25-JUN-1998; 98US-0090688P.  
 PR 25-JUN-1998; 98US-0090690P.



PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 22-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 28-MAY-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 11-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089533P.  
PR 18-JUN-1998; 98US-0089539P.  
PR 19-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090684P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 26-JUN-1998; 98US-0091159P.  
PR 01-JUL-1998; 98US-0091154P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091828P.  
PR 02-JUL-1998; 98US-0091832P.  
PR 02-JUL-1998; 98US-0094006P.  
PR 24-JUL-1998; 98US-0095282P.  
PR 04-AUG-1998; 98US-0095988P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 10-AUG-1998; 98US-0096157P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098016P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-0101930.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101058P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101942P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.

```

PR 01-OCT-1998; 98US-0102687P.
Alignment Scores:
Pred. No.: 147
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.63%
DB: 7
Gaps: 0

US-09-690-825-34 (1-142) x ACC87608 (1-1734)
QY 1 MetGlyAlaProThrLeuProPro 8
Db 1534 ATGGGAGCCCCACACTCCCTCCT 1557

RESULT 52
ACC86994
ID ACC86994 standard; cDNA; 1734 BP.
AC ACC86994;
XX
XX
DT 05-AUG-2003 (first entry)
XX
DE Human secreted polypeptide PRO1411-encoding cDNA, SEQ ID NO:201.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnary; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
FN US2003036159-A1.
XX
PD 20-FEB-2003.
XX
XX
PF 02-JUL-2002; 2002US-00188773.
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 21-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 11-DEC-1997; 97US-0066772P.
PR 12-DEC-1997; 97US-0069335P.
PR 17-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 10-MAR-1998; 98US-0076450P.
PR 11-MAR-1998; 98US-0077632P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083553P.
PR 05-MAY-1998; 98US-0083568P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 22-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 28-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087603P.
PR 02-JUN-1998; 98US-0087753P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088328P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 11-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089588P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.

```



PR 21-APR-1998; 98US-0082563P.  
 PR 22-APR-1998; 98US-0082704P.  
 PR 22-APR-1998; 98US-0082797P.  
 PR 28-APR-1998; 98US-0083322P.  
 PR 29-APR-1998; 98US-0083495P.  
 PR 29-APR-1998; 98US-0083496P.  
 PR 29-APR-1998; 98US-0083499P.  
 PR 29-APR-1998; 98US-0083559P.  
 PR 05-MAY-1998; 98US-0084366P.  
 PR 06-MAY-1998; 98US-0084414P.  
 PR 07-MAY-1998; 98US-0084639P.  
 PR 07-MAY-1998; 98US-0084640P.  
 PR 07-MAY-1998; 98US-0084643P.  
 PR 15-MAY-1998; 98US-0085579P.  
 PR 15-MAY-1998; 98US-0085580P.  
 PR 15-MAY-1998; 98US-0085582P.  
 PR 15-MAY-1998; 98US-0085700P.  
 PR 18-MAY-1998; 98US-0086023P.  
 PR 22-MAY-1998; 98US-0086392P.  
 PR 22-MAY-1998; 98US-0086486P.  
 PR 28-MAY-1998; 98US-0087098P.  
 PR 28-MAY-1998; 98US-0087208P.  
 PR 02-JUN-1998; 98US-0087609P.  
 PR 02-JUN-1998; 98US-0087759P.  
 PR 03-JUN-1998; 98US-0087827P.  
 PR 04-JUN-1998; 98US-0088025P.  
 PR 04-JUN-1998; 98US-0088028P.  
 PR 04-JUN-1998; 98US-0088039P.  
 PR 04-JUN-1998; 98US-0088033P.  
 PR 04-JUN-1998; 98US-0088326P.  
 PR 05-JUN-1998; 98US-0088167P.  
 PR 05-JUN-1998; 98US-0088202P.  
 PR 05-JUN-1998; 98US-0088212P.  
 PR 05-JUN-1998; 98US-0088217P.  
 PR 09-JUN-1998; 98US-0088655P.  
 PR 10-JUN-1998; 98US-0088722P.  
 PR 10-JUN-1998; 98US-0088738P.  
 PR 10-JUN-1998; 98US-0088740P.  
 PR 10-JUN-1998; 98US-0088811P.  
 PR 10-JUN-1998; 98US-0088824P.  
 PR 10-JUN-1998; 98US-0088825P.  
 PR 10-JUN-1998; 98US-0088826P.  
 PR 11-JUN-1998; 98US-0088861P.  
 PR 11-JUN-1998; 98US-0088863P.  
 PR 11-JUN-1998; 98US-0088876P.  
 PR 12-JUN-1998; 98US-0089090P.  
 PR 12-JUN-1998; 98US-0089105P.  
 PR 16-JUN-1998; 98US-0089512P.  
 PR 16-JUN-1998; 98US-0089514P.  
 PR 17-JUN-1998; 98US-0089538P.  
 PR 17-JUN-1998; 98US-0089598P.  
 PR 17-JUN-1998; 98US-0089653P.  
 PR 18-JUN-1998; 98US-0089908P.  
 PR 19-JUN-1998; 98US-0089952P.  
 PR 22-JUN-1998; 98US-0090246P.  
 PR 22-JUN-1998; 98US-0090252P.  
 PR 22-JUN-1998; 98US-0090254P.  
 PR 24-JUN-1998; 98US-0090429P.  
 PR 24-JUN-1998; 98US-0090435P.  
 PR 24-JUN-1998; 98US-0090444P.  
 PR 24-JUN-1998; 98US-0090461P.  
 PR 24-JUN-1998; 98US-0090535P.  
 PR 24-JUN-1998; 98US-0090540P.  
 PR 25-JUN-1998; 98US-0090676P.  
 PR 25-JUN-1998; 98US-0090678P.  
 PR 25-JUN-1998; 98US-0090688P.  
 PR 25-JUN-1998; 98US-0090690P.  
 PR 25-JUN-1998; 98US-0090694P.  
 PR 25-JUN-1998; 98US-0090695P.  
 PR 26-JUN-1998; 98US-0090696P.  
 PR 26-JUN-1998; 98US-00905413.  
 PR 26-JUN-1998; 98US-0090862P.  
 PR 26-JUN-1998; 98US-0090863P.

PR 26-JUN-1998; 98US-0091010P.  
 PR 01-JUL-1998; 98US-0091359P.  
 PR 01-JUL-1998; 98US-0091544P.  
 PR 02-JUL-1998; 98US-0091478P.  
 PR 02-JUL-1998; 98US-0091486P.  
 PR 02-JUL-1998; 98US-0091626P.  
 PR 02-JUL-1998; 98US-0091628P.  
 PR 02-JUL-1998; 98US-0091633P.  
 PR 24-JUL-1998; 98US-0094006P.  
 PR 04-AUG-1998; 98US-0095282P.  
 PR 10-AUG-1998; 98US-0095998P.  
 PR 10-AUG-1998; 98US-0096012P.  
 PR 17-AUG-1998; 98US-0096757P.  
 PR 17-AUG-1998; 98US-0096766P.  
 PR 17-AUG-1998; 98US-0096867P.  
 PR 17-AUG-1998; 98US-0096891P.  
 PR 17-AUG-1998; 98US-0096897P.  
 PR 18-AUG-1998; 98US-0096949P.  
 PR 18-AUG-1998; 98US-0096959P.  
 PR 18-AUG-1998; 98US-0097023P.  
 PR 26-AUG-1998; 98US-0097953P.  
 PR 26-AUG-1998; 98US-0097954P.  
 PR 26-AUG-1998; 98US-0097955P.  
 PR 26-AUG-1998; 98US-0097971P.  
 PR 26-AUG-1998; 98US-0097974P.  
 PR 26-AUG-1998; 98US-0098014P.  
 PR 01-SEP-1998; 98US-0098716P.  
 PR 01-SEP-1998; 98US-0098723P.  
 PR 02-SEP-1998; 98US-0098803P.  
 PR 02-SEP-1998; 98US-0098821P.  
 PR 02-SEP-1998; 98US-0098843P.  
 PR 09-SEP-1998; 98US-0099602P.  
 PR 10-SEP-1998; 98US-0099741P.  
 PR 10-SEP-1998; 98US-0099754P.  
 PR 10-SEP-1998; 98US-0099763P.  
 PR 10-SEP-1998; 98US-0099812P.  
 PR 15-SEP-1998; 98US-0100388P.  
 PR 16-SEP-1998; 98US-0100662P.  
 PR 16-SEP-1998; 98US-0100664P.  
 PR 16-SEP-1998; 98US-0101751P.  
 PR 16-SEP-1998; 98US-0101751P.  
 PR 17-SEP-1998; 98US-0100683P.  
 PR 17-SEP-1998; 98US-0100684P.  
 PR 17-SEP-1998; 98US-0100919P.  
 PR 17-SEP-1998; 98US-0100930P.  
 PR 18-SEP-1998; 98US-0100849P.  
 PR 18-SEP-1998; 98US-0101014P.  
 PR 18-SEP-1998; 98US-0101068P.  
 PR 23-SEP-1998; 98US-0101471P.  
 PR 23-SEP-1998; 98US-0101472P.  
 PR 23-SEP-1998; 98US-0101475P.  
 PR 23-SEP-1998; 98US-0101477P.  
 PR 24-SEP-1998; 98US-0101738P.  
 PR 24-SEP-1998; 98US-0101739P.  
 PR 24-SEP-1998; 98US-0101743P.  
 PR 24-SEP-1998; 98US-0101922P.  
 PR 25-SEP-1998; 98US-0101786P.  
 PR 29-SEP-1998; 98US-0102207P.  
 PR 29-SEP-1998; 98US-0102240P.  
 PR 29-SEP-1998; 98US-0102330P.  
 PR 29-SEP-1998; 98US-0102331P.  
 PR 30-SEP-1998; 98US-0102487P.  
 PR 30-SEP-1998; 98US-0102570P.  
 PR 30-SEP-1998; 98US-0102571P.  
 PR 01-OCT-1998; 98US-0102684P.  
 PR 01-OCT-1998; 98US-0102687P.  
 PR 02-OCT-1998; 98US-0102965P.  
 PR 06-OCT-1998; 98US-0103258P.  
 PR 06-OCT-1998; 98US-0103449P.  
 PR 07-OCT-1998; 98US-00968978.  
 PR 07-OCT-1998; 98US-00968978.

Alignment Scores:  
 Pred. NO.:

147

Length:

1734

Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 7 Gaps: 0

US-09-690-825-34 (1-142) x ACD04167 (1-1734)

QY 1 MetGlyAlaProThreuProPro 8  
DB 1534 ATGGGAGCCCCACACTCCCTCCT 1557

RESULT 54  
AC69498  
ID ACA69498 standard; cDNA; 1734 BP.  
XX  
AC ACA69498;  
XX  
DT 27-JUN-2003 (first entry)  
XX  
DE cDNA encoding human PRO polypeptide #101.  
KW Human; PRO polypeptide; secreted and transmembrane protein; tumour;  
KW chromosome mapping; gene mapping; cytostatic; gene therapy; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN US2003032113-A1.  
XX  
PD 13-FEB-2003.  
XX  
PF 20-JUN-2002; 2002US-00176911.  
XX  
PR 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 27-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080337P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082588P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.

PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083435P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086332P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090439P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.





PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083497P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084633P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0088027P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088367P.  
PR 05-JUN-1998; 98US-0088402P.  
PR 05-JUN-1998; 98US-0088512P.  
PR 09-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.

PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095958P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096876P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 03-SEP-1998; 98US-0098902P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 15-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100642P.  
PR 16-SEP-1998; 98US-0100644P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101942P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-00168978.

Alignment Scores:  
Pred. No.: 147

Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 5.63%

Length: 1734  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0

```
DB: 7 Gaps: 0
US-09-690-825-34 (1-142) x ACA90343 (1-1734)
QY 1 MetGlyAlaProThrLeuProPro 8
Dd 1534 ATGGAGGCCGCCACACTCCCTCCT 1557

RESULT 56
ACCR9450
ID ACC89450 standard; cDNA; 1734 BP.
AC
AC ACC89450;
DI 11-AUG-2003 (first entry)
DE Human secreted polypeptide PRO1411-encoding cDNA, SEQ ID NO:201.
XX
XX Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnery; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
XX US2003027264-A1.
PD 06-FEB-2003.
XX
XX 18-JUN-2002; 2002US-00174579.
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 31-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082586P.
PR 21-APR-1998; 98US-0082589P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087753P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 26-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
```





QY 1 MetGlyAlaProThrIleuProPro 8  
Db 1534 ATGGGAGCCCCACACTCCCTCT 1557  
RESULT 58  
ACA93883  
ID ACA93883 standard; cDNA; 1734 BP.  
XX  
AC ACA93883;  
XX  
XX 18-JUL-2003 (first entry)  
XX  
DE Human secreted/transmembrane protein (PRO) cDNA #101.  
XX  
XX Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;  
KW proliferation; differentiation; chondrocyte cell; TNF-alpha;  
KW tumour necrosis factor-alpha; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN US2003036149-A1.  
XX  
PD 20-FEB-2003.  
XX  
XX 02-JUL-2002; 2002US-00187746.  
XX  
PR 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0064250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 26-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090446P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.

```

PR 10-AUG-1998; 98US-0095938P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98MO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 18-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 23-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Alignment Scores:
Pred. No.: 147
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.63%
DB: 7

Length: 1734
Matches: 8
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-690-825-34 (1-142) x ACA93883 (1-1734)
QY 1 MetGlyAlaProThrLeuProPro 8
|||||

Db 1534 ATGGAGCCCCCACACTCCCTCT 1557
RESULT 60

Db 1534 ATGGAGCCCCCACACTCCCTCT 1557
RESULT 59
ACA66918
ID ACA66918 standard; cDNA; 1734 BP.
XX ACA66918;
AC ACA66918;
XX 23-JUN-2003 (first entry)
XX cDNA encoding human PRO polypeptide #78.
DE Human; PRO polypeptide; secreted and transmembrane protein;
KW anti-PRO antibody; diagnostic assay; gene expression; tumour; cytostatic;
KW gene; ss.
XX Homo sapiens.
XX US2003036635-A1.
XX 20-FEB-2003.
XX 28-AUG-2002; 2002US-00230163.
XX 25-JUL-2000; 2000US-0220638P.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX (GETH ) GENENTECH INC.
XX Baker KP, Deenoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX WPI; 2003-342045/32.
DR P-PSDB; ABU80816.
XX One hundred and twenty two nucleic acids encoding PRO polypeptides,
PT useful for the manufacture of a medicament for diagnosing or treating
PT tumor.
XX Claim 2; Fig 155; 314pp; English.
XX The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The PRO
CC polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides and polynucleotides are useful for preparing a medicament
CC useful in the diagnosis and treatment of tumours. Anti-PRO antibodies are
CC useful in diagnostic assays for PRO, by detecting its expression in
CC specific cells, tissues or serum, and for affinity purification of PRO
CC from recombinant cell culture or natural sources. ACA66841-ACA66962
CC represent cDNA sequences encoding the human PRO polypeptides of the
CC invention. Note: The sequence data for this patent was obtained in
CC electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/psaidsentry.html
XX
SQ Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 147
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.63%
DB: 7

Length: 1734
Matches: 8
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-690-825-34 (1-142) x ACA66918 (1-1734)
QY 1 MetGlyAlaProThrLeuProPro 8
|||||

Db 1534 ATGGAGCCCCCACACTCCCTCT 1557
RESULT 60

```

ACD15276  
ID ACD15276 standard; cDNA; 1734 BP.  
XX  
AC ACD15276;  
XX  
DT 17-AUG-2003 (first entry)  
XX  
DE Human secreted/transmembrane protein (PRO) cDNA #101.  
XX  
KW Human; gene; ss; secreted and transmembrane protein; PRO; TNF-alpha;  
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;  
KW tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;  
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.  
XX  
OS Homo sapiens.  
XX  
PN US2003044923-A1.  
XX  
PD 06-MAR-2003.  
XX  
PF 24-JUN-2002; 2002US-00179522.  
XX  
PR 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069970P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083552P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088126P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088555P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090451P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.

```

PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096881P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

```

```

Alignment Scores: 147 1734
Pred. No.: 8.00
Score: 100.00%
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.63%
DB: 7
Length: 1734
Matches: 8
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

```

US-09-690-825-34 (1-142) x ACD15276 (1-1734)

```

Oy 1 MetClyAlaProThrLeuProPro 8
Db 1534 ATGGGAGCCCCACACTCCCTCT 1557

```

RESULT 61

ACD08863

ID ACD08863 standard; cDNA; 1734 BP.

```

XX ACD08863;
AC 09-AUG-2003 (first entry)
DT 09-AUG-2003 (first entry)
DE Human secreted/transmembrane protein (PRO) cDNA #101.
XX Human; Gene: ss; secreted and transmembrane protein; PRO; TNF-alpha;
XX tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
XX tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;
XX prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX Homo sapiens.
XX US2003040062-A1.
PD 27-FEB-2003.
XX 25-JUN-2002; 2002US-00180545.
XX 18-SEP-1997; 97US-0059263P.
XX 18-SEP-1997; 97US-0059266P.
XX 17-OCT-1997; 97US-0062250P.
XX 21-OCT-1997; 97US-0063486P.
XX 24-OCT-1997; 97US-0063120P.
XX 24-OCT-1997; 97US-0063121P.
XX 28-OCT-1997; 97US-0063540P.
XX 28-OCT-1997; 97US-0063541P.
XX 28-OCT-1997; 97US-0063544P.
XX 28-OCT-1997; 97US-0063564P.
XX 29-OCT-1997; 97US-0063734P.
XX 31-OCT-1997; 97US-0063870P.
XX 31-OCT-1997; 97US-0064103P.
XX 13-NOV-1997; 97US-0065311P.
XX 21-NOV-1997; 97US-0066120P.
XX 24-NOV-1997; 97US-0066466P.
XX 11-DEC-1997; 97US-0066772P.
XX 12-DEC-1997; 97US-0069425P.
XX 17-DEC-1997; 97US-0069870P.
XX 18-DEC-1997; 97US-0068017P.
XX 10-MAR-1998; 98US-0077450P.
XX 11-MAR-1998; 98US-0077632P.
XX 11-MAR-1998; 98US-0077649P.
XX 20-MAR-1998; 98US-0078866P.
XX 20-MAR-1998; 98US-0078939P.
XX 27-MAR-1998; 98US-0079664P.
XX 27-MAR-1998; 98US-0079786P.
XX 31-MAR-1998; 98US-0080107P.
XX 31-MAR-1998; 98US-0080194P.
XX 01-APR-1998; 98US-0080327P.
XX 01-APR-1998; 98US-0080333P.
XX 08-APR-1998; 98US-0081049P.
XX 08-APR-1998; 98US-0081070P.
XX 09-APR-1998; 98US-0081195P.
XX 15-APR-1998; 98US-0081838P.
XX 21-APR-1998; 98US-0082568P.
XX 21-APR-1998; 98US-0082569P.
XX 22-APR-1998; 98US-0082704P.
XX 22-APR-1998; 98US-0082797P.
XX 28-APR-1998; 98US-0083322P.
XX 29-APR-1998; 98US-0083495P.
XX 29-APR-1998; 98US-0083496P.
XX 29-APR-1998; 98US-0083499P.
XX 29-APR-1998; 98US-0083559P.
XX 05-MAY-1998; 98US-0084366P.
XX 06-MAY-1998; 98US-0084444P.
XX 07-MAY-1998; 98US-0084639P.
XX 07-MAY-1998; 98US-0084640P.
XX 07-MAY-1998; 98US-0084643P.
XX 15-MAY-1998; 98US-0085579P.
XX 15-MAY-1998; 98US-0085580P.
XX 15-MAY-1998; 98US-0085582P.

```



PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 26-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-0090699P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.

PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0098902P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-00999812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.

## Alignment Scores:

Pred. No.: 147  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 5.63%  
DB: 7  
Length: 1734  
Matches: 8  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-09-690-825-34 (1-142) x ACD08863 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8  
|||  
Db 1534 ATGGAGGCCCCACACTCCCTCCT 1557

## RESULT 62

ACC96783  
ID ACC96783 standard; cDNA; 1734 BP.  
XX  
AC ACC96783;

XX 10-SEP-2003 (first entry)  
DT Human secreted polypeptide PRO1411-encoding cDNA, SEQ ID NO:201.  
DE  
XX  
XX Human; PRO; secreted protein; transmembrane protein;  
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;  
KW chondrocyte; proliferation; differentiation; cartilage disorder;  
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;  
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;  
KW liver; drug screening; transgenic animal; genetic analysis;  
XX antiarthritic; vulnary; gene therapy; gene; ss.  
OS Homo sapiens.  
XX  
XX US2003040056-A1.  
PN  
XX  
XX 27-FEB-2003.  
PD  
XX  
XX 21-JUN-2002; 2002US-00176916.  
PF  
XX 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066456P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069333P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078866P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079564P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081043P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082737P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086466P.  
PR 22-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087659P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090698P.  
PR 26-JUN-1998; 98US-00905413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.

```
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096849P.
PR 18-AUG-1998; 98US-0096859P.
PR 18-AUG-1998; 98US-0097020P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0098602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 25-SEP-1998; 98US-0102207P.
PR 25-SEP-1998; 98US-0102240P.
PR 25-SEP-1998; 98US-0102330P.
PR 25-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096849P.
PR 18-AUG-1998; 98US-0096859P.
PR 18-AUG-1998; 98US-0097020P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0098602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 25-SEP-1998; 98US-0102207P.
PR 25-SEP-1998; 98US-0102240P.
PR 25-SEP-1998; 98US-0102330P.
PR 25-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

XX DE Human secreted polypeptide PRO1411-encoding cDNA, SEQ ID NO:201.
XX KW Human; PRO; secreted protein; transmembrane protein;
XX KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
XX KW chondrocyte; proliferation; differentiation; cartilage disorder;
XX KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
XX KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
XX KW liver; drug screening; transgenic animal; genetic analysis;
XX KW antiarthritic; vulnery; gene therapy; gene; ss.
XX OS Homo sapiens.
XX FN US2003044926-A1.
XX PD 06-MAR-2003.
XX PP 26-JUN-2002; 2002US-00183015.
XX PR 18-SEP-1997; 97US-0059263P.
XX PR 18-SEP-1997; 97US-0059266P.
XX PR 17-OCT-1997; 97US-0062250P.
XX PR 21-OCT-1997; 97US-0063486P.
XX PR 24-OCT-1997; 97US-0063120P.
XX PR 24-OCT-1997; 97US-0063121P.
XX PR 28-OCT-1997; 97US-0063540P.
XX PR 28-OCT-1997; 97US-0063541P.
XX PR 28-OCT-1997; 97US-0063544P.
XX PR 28-OCT-1997; 97US-0063564P.
XX PR 29-OCT-1997; 97US-0063734P.
XX PR 31-OCT-1997; 97US-0063870P.
XX PR 31-OCT-1997; 97US-0064103P.
XX PR 13-NOV-1997; 97US-0065311P.
XX PR 21-NOV-1997; 97US-0066120P.
XX PR 24-NOV-1997; 97US-0066466P.
XX PR 11-DEC-1997; 97US-0066772P.
XX PR 12-DEC-1997; 97US-0069335P.
XX PR 12-DEC-1997; 97US-0069425P.
XX PR 18-DEC-1997; 97US-0069870P.
XX PR 18-DEC-1997; 97US-0068017P.
XX PR 10-MAR-1998; 98US-0077450P.
XX PR 11-MAR-1998; 98US-0077632P.
XX PR 11-MAR-1998; 98US-0077649P.
XX PR 20-MAR-1998; 98US-0078866P.
XX PR 20-MAR-1998; 98US-0078939P.
XX PR 27-MAR-1998; 98US-0079664P.
XX PR 27-MAR-1998; 98US-0079786P.
XX PR 31-MAR-1998; 98US-0080107P.
XX PR 31-MAR-1998; 98US-0080194P.
XX PR 01-APR-1998; 98US-0080327P.
XX PR 01-APR-1998; 98US-0080333P.
XX PR 08-APR-1998; 98US-0081049P.
XX PR 08-APR-1998; 98US-0081070P.
XX PR 09-APR-1998; 98US-0081195P.
XX PR 15-APR-1998; 98US-0081838P.
XX PR 21-APR-1998; 98US-0082568P.
XX PR 21-APR-1998; 98US-0082569P.
XX PR 22-APR-1998; 98US-0082704P.
XX PR 22-APR-1998; 98US-0082797P.
XX PR 28-APR-1998; 98US-0083322P.
XX PR 28-APR-1998; 98US-0083485P.
XX PR 29-APR-1998; 98US-0083496P.
XX PR 29-APR-1998; 98US-0083499P.
XX PR 29-APR-1998; 98US-0083559P.
XX PR 05-MAY-1998; 98US-0084366P.
XX PR 06-MAY-1998; 98US-0084414P.
XX PR 07-MAY-1998; 98US-0084639P.
XX PR 07-MAY-1998; 98US-0084640P.
XX PR 07-MAY-1998; 98US-0084643P.
XX PR 15-MAY-1998; 98US-0085579P.
XX PR 15-MAY-1998; 98US-0085580P.
XX PR 15-MAY-1998; 98US-0085582P.
XX PR 15-MAY-1998; 98US-0085700P.
```

```
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087599P.
PR 02-JUN-1998; 98US-0087599P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088033P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088212P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088835P.
PR 10-JUN-1998; 98US-0088862P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090246P.
PR 24-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 02-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 04-AUG-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095988P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.

PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098715P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 25-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Alignment Scores:
Pred. No.: 147
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.63%
DB: 7

Length: 1734
Matches: 8
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-690-825-34 (1-142) x ACF15504 (1-1734)
QY 1 MetGlyAlaProThrLeuProPro 8
Db 1534 ATGGAGGCCGCCACACTCCCTCT 1557

RESULT 64
ACD42388
ID ACD42388 standard; cDNA; 1734 BP.
XX ACD42388;
AC ACD42388;
XX
DT 05-SEP-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO141 cDNA.
```

XX Human; secreted and transmembrane protein; PRO: anti-diabetic;  
KW ophthalmological; cytostatic; immunostimulant; gene therapy;  
KW vascular endothelial growth factor inhibitor; hypertrophy of adult heart;  
KW protein secretion disorder; pancreas disorder; diabetes;  
KW vascular permeability; retinal neuron cell survival; retinal disorder;  
KW immune response; inflammation; mononuclear cell infiltration;  
KW eosinophil infiltration; apoptosis; neoplastic growth; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN US2003040014-A1.  
XX  
PD 27-FEB-2003.  
XX  
PF 01-FEB-2002; 2002US-00066269.  
XX  
PR 26-AUG-1997; 97US-0056974P.  
PR 17-SEP-1997; 97US-0059115P.  
PR 18-SEP-1997; 97US-0059263P.  
PR 19-SEP-1997; 97US-0059588P.  
PR 17-OCT-1997; 97US-0062285P.  
PR 24-OCT-1997; 97US-0062815P.  
PR 24-OCT-1997; 97US-0063082P.  
PR 24-OCT-1997; 97US-0063329P.  
PR 29-OCT-1997; 97US-0063733P.  
PR 21-NOV-1997; 97US-0066364P.  
PR 25-NOV-1997; 97US-0066840P.  
PR 16-DEC-1997; 97US-0069694P.  
PR 09-FEB-1998; 98US-0074086P.  
PR 09-FEB-1998; 98US-0074092P.  
PR 09-FEB-1998; 98US-0079294P.  
PR 08-APR-1998; 98US-0081049P.  
PR 14-JUL-1998; 98WO-US014552.  
PR 18-AUG-1998; 98US-0095998P.  
PR 18-AUG-1998; 98US-0097000P.  
PR 09-SEP-1998; 98US-0099601P.  
PR 10-SEP-1998; 98US-0099803P.  
PR 10-SEP-1998; 98US-0099811P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100858P.  
PR 17-SEP-1998; 98WO-US019437.  
PR 24-SEP-1998; 98US-0101922P.  
PR 28-OCT-1998; 98US-0106032P.  
PR 20-NOV-1998; 98US-0109304P.  
PR 20-NOV-1998; 98WO-US024855.  
PR 23-NOV-1998; 98WO-US025190.  
PR 01-DEC-1998; 98WO-US025108.  
PR 08-MAR-1999; 99WO-US005028.  
PR 23-MAR-1999; 99US-0125778P.  
PR 02-JUN-1999; 99WO-US012252.  
PR 15-JUN-1999; 99US-0139695P.  
PR 20-JUL-1999; 99US-0145070P.  
PR 26-JUL-1999; 99US-0145698P.  
PR 17-AUG-1999; 99US-0149396P.  
PR 01-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 30-NOV-1999; 99WO-US028313.  
PR 01-DEC-1999; 99WO-US028301.  
PR 02-DEC-1999; 99WO-US028565.  
PR 07-DEC-1999; 99US-0169495P.  
PR 20-DEC-1999; 99WO-US030999.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 09-MAR-2000; 2000WO-US006471.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 15-MAY-2000; 2000WO-US013358.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 28-FEB-2001; 2001WO-US008520.  
PR 30-MAY-2001; 2001WO-US017443.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 15-NOV-2001; 2001US-00002796.  
XX  
FA (GETH ) GENENTECH INC.  
XX  
XX Ashkenazi AJ, Baker KP, Botstein DA, Desnoyers L, Eaton DL,  
PI Ferrara N, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;  
PI Godowski PJ, Gurney AL, Kljavin LJ, Mather JP, Napier MA, Pan J;  
PI Paoni NF, Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM;  
PI Wood WI, Zhang Z;  
XX  
DR WPI; 2003-503396/47.  
XX P-PSDB; ABO25176.  
XX  
PT New secreted and transmembrane PRO polypeptides, useful for treating  
PT diabetes, retinal disorders and stimulating an immune response.  
XX  
XX Claim 2; Fig 53; 254pp; English.  
XX  
CC The invention describes an isolated polypeptide (1) having at least 80 %  
CC amino acid sequence identity to 30 secreted and transmembrane  
CC polypeptides. PRO polypeptides are also useful for stimulating  
CC hypertrophy of adult heart, for inhibiting vascular endothelial growth  
CC factor stimulated proliferation of endothelial cells, stimulating  
CC proliferation of stimulated T-lymphocytes and for inducing proliferation  
CC of PD12 pancreatic ductal cells and are thus useful in the treatment of  
CC disorders which involve protein secretion by the pancreas, including  
CC diabetes. PRO polypeptides are useful for inducing vascular permeability  
CC and in enhancing survival of retinal neurons cells and are thus useful  
CC for the treatment of retinal disorders. PRO polypeptides are also useful  
CC for stimulating an immune response and inducing inflammation by inducing  
CC mononuclear cell and eosinophil infiltration at the site of infection of  
CC an animal. The PRO polypeptides are further useful for inducing apoptosis  
CC in endothelial cells for inhibiting neoplastic growth. This sequence  
CC encodes a novel human secreted and transmembrane PRO polypeptide  
XX  
SQ Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 147 Length: 1734  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 7 Gaps: 0  
US-09-690-825-34 (1-142) x ACD42388 (1-1734)  
QY 1 MetGlyAlaProThrLeuProPro 8  
DB 1534 ATGGAGGCCCCACACTCCCTCCT 1557  
RESULT 65  
ACD68670  
ID ACD68670 standard; cDNA; 1734 BP.  
XX

```

AC ACDC68670;
AD 17-SEP-2003 (first entry)
DE Novel human secreted and transmembrane protein PRO1411 cDNA.
DX
KW Human; secreted and transmembrane protein; PRO; cytostatic;
KW antiarthritic; osteopathic; gene therapy; TNF-Agonist-Alpha;
KW chondrocyte stimulator; pericyte stimulator; fibroblast modulator;
KW pharmaceutical; diagnostic; biosensor; bioreactor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; bone disorder; cartilage disorder; sports injury;
KW arthritis; wound; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003045687-A1.
XX
PD 06-MAR-2003.
XX
PF 12-AUG-2002; 2002US-00218631.
XX
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ,
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
DR WPI; 2003-512315/48.
DR P-PSDB; ABO33782.
XX
XX New genes, and its encoded secreted and transmembrane polypeptides,
PT useful for stimulating Tumor Necrosis Factor alpha, or chondrocyte or
PT pericyte proliferation, especially for treating lung tumors, arthritis or
PT wounds in a mammal.
XX
PS Claim 2; Fig 155; 314pp; English.
XX
CC The invention describes an isolated nucleic acid molecule comprising a
CC sequence with at least 80% identity to: (a) a nucleotide encoding any of
CC 122 PRO (secreted and transmembrane) polypeptides whose sequences are
CC fully defined in the specification; or (b) any of 122 nucleotide
CC sequences having e.g. 4834, 2504 or 1759 bp fully defined in the
CC specification; or the full length coding sequence of any these 122
CC nucleotide sequences. The PRO polypeptides or polynucleotides are useful
CC as pharmaceuticals, diagnostics, biosensors or bioreactors. These are
CC particularly useful for detecting tumors (e.g. lung tumour, colon
CC tumour, breast tumour, prostate tumour, rectal tumour, or liver tumour)
CC in a mammal, for stimulating the release of TNF-alpha from human blood,
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells, for stimulating proliferation of pericyte cells, or for modulating
CC normal human dermal fibroblast proliferation. The PRO nucleic acid or
CC polypeptide is also useful for treating tumors or various bone and/or
CC cartilage disorders (e.g. sports injuries or arthritis), or wounds. The
CC PRO polypeptides are useful in drug screening, particularly as targets
CC for therapeutic intervention in these diseases, and in the diagnostic
CC determination of the presence of these diseases. The PRO polypeptides are
CC also useful as molecular weight markers, or for chromosome
CC identification. The PRO genes are useful as hybridisation probes, or for
CC screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may
CC also be used in gene therapy, particularly for replacing a defective
CC gene. This sequence encodes a novel human secreted and transmembrane PRO
CC polypeptide
XX
SQ Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.: 147 Length: 1734

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.63% Indels: 0  
 DB: 7 Gaps: 0

US-09-690-825-34 (1-142) x ACD68670 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8  
 DB 1534 ATGGGAGCCCCCACACTCCCTCT 1557

RESULT 66

ACA72871

ID ACA72871 standard; cDNA; 1734 BP.

XX

AC ACA72871;

XX

DT 02-JUL-2003 (first entry)

XX

DE Human PRO polynucleotide #101.

XX

KW Human; PRO; Gene; ss; secreted polypeptide; transmembrane polypeptide;  
 KW tumour necrosis factor alpha; TNF-alpha; chondrocyte cell; tumour;  
 KW cytostatic.

XX

OS Homo sapiens.

XX

PN US2003036140-A1.

XX

PD 20-FEB-2003.

XX

PF 01-JUL-2002; 2002US-00187588.

XX

XX 26-JUN-1998; 98US-00105413.

PR 16-SEP-1998; 98WO-US019330.

PR 07-OCT-1998; 98US-00168978.

PR 07-OCT-1998; 98WO-US021141.

PR 06-NOV-1998; 98US-00187388.

PR 01-DEC-1998; 98WO-US025108.

PR 07-DEC-1998; 98US-00202054.

PR 03-MAR-1999; 99US-00254311.

PR 08-MAR-1999; 99WO-US005028.

PR 14-MAY-1999; 99US-00311832.

PR 14-MAY-1999; 99WO-US010733.

PR 25-AUG-1999; 99US-00380137.

PR 25-AUG-1999; 99US-00380138.

PR 25-AUG-1999; 99US-00380139.

PR 25-AUG-1999; 99US-00380142.

PR 01-SEP-1999; 99WO-US020111.

PR 15-SEP-1999; 99WO-US021090.

PR 18-OCT-1999; 99US-00403297.

PR 12-NOV-1999; 99US-00423844.

PR 01-DEC-1999; 99WO-US028301.

PR 02-DEC-1999; 99WO-US028551.

PR 30-DEC-1999; 99WO-US031274.

PR 05-JAN-2000; 2000WO-US000219.

PR 18-FEB-2000; 2000WO-US004341.

PR 18-FEB-2000; 2000WO-US004342.

PR 22-FEB-2000; 2000WO-US004414.

PR 24-FEB-2000; 2000WO-US005004.

PR 01-MAR-2000; 2000WO-US005601.

PR 02-MAR-2000; 2000WO-US005841.

PR 15-MAR-2000; 2000WO-US006884.

PR 30-MAR-2000; 2000WO-US008439.

PR 17-MAY-2000; 2000WO-US013705.

PR 22-MAY-2000; 2000WO-US014042.

PR 30-MAY-2000; 2000WO-US014941.

PR 02-JUN-2000; 2000WO-US015264.

PR 28-JUL-2000; 2000WO-US020710.

PR 22-AUG-2000; 2000US-0064848.

PR 24-AUG-2000; 2000WO-US023328.

PR 18-SEP-2000; 2000WO-US026410.

PR 18-SEP-2000; 2000US-00665350.

PR 08-NOV-2000; 2000US-00709238.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 22-MAR-2001; 2001US-00816744.  
PR 10-MAY-2001; 2001US-00854208.  
PR 10-MAY-2001; 2001US-00854280.  
PR 25-MAY-2001; 2001US-00866028.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 30-JUL-2001; 2001US-00918588.  
PR 06-AUG-2001; 2001US-00924419.  
PR 13-AUG-2001; 2001US-00929404.  
PR 16-AUG-2001; 2001US-00931836.  
PR 28-AUG-2001; 2001US-00941992.  
PR 29-AUG-2001; 2001WO-US027099.  
PR 04-SEP-2001; 2001US-00946374.  
PR 15-JAN-2002; 2002US-00025856.  
XX  
PA (GETH ) GENENTECH INC.  
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;  
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;  
XX  
DR WPI; 2003-332028/31.  
DR P-PSDB; AB085678.  
XX  
PT Three hundred and five nucleic acids encoding PRO polypeptides, useful  
PT for the manufacture of a medicament for diagnosing or treating tumor.  
XX  
XX Claim 2; Fig 201; 707pp; English.  
XX  
CC The invention relates to human PRO polypeptides (secreted and  
CC transmembrane polypeptides) and the PRO polynucleotides encoding them.  
CC The invention also relates to a method for stimulating the release of  
CC tumour necrosis factor alpha (TNF-alpha) from human blood by contacting  
CC the blood with a sequence of the invention, a method for stimulating the  
CC proliferation or differentiation of chondrocyte cells by contacting the  
CC cells with a PRO polypeptide and a method for detecting the presence of a  
CC tumour in a mammal. The polypeptides and polynucleotides are useful for  
CC the manufacture of a medicament for diagnosing or treating a tumour in a  
CC mammal. Sequences ACA72771-ACA73075 represent human PRO polynucleotides  
CC of the invention. Note: The sequence data for this patent is also  
CC available in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 147 Length: 1734  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 7 Gaps: 0

US-09-690-825-34 (1-142) x ACA72871 (1-1734)  
Qy 1 MetGlyAlaProThrLeuProPro 8  
Db 1534 ATGGGAGCCCCCACACTCCCTCT 1557

RESULT 67  
ACD03043  
ID ACD03043 standard; cDNA; 1734 BP.  
XX  
AC ACD03043;

XX 01-AUG-2003 (first entry)  
XX Novel human secreted and transmembrane protein PRO1411 cDNA.  
DE  
XX Human; secreted and transmembrane protein; PRO; cytostatic; gene therapy;  
KW chondrocyte stimulator; tumour; adrenal tumour; lung tumour;  
KW colon tumour; breast tumour; prostate tumour; rectal tumour;  
KW cervical tumour; liver tumour; TNF-alpha release;  
KW tumour necrosis factor alpha release; chondrocyte cell proliferation;  
KW chondrocyte cell differentiation; pharmaceutical; diagnostic; biosensor;  
KW bioreactor; gene; ss.  
XX Homo sapiens.  
XX US2003013153-A1.  
XX 16-JAN-2003.  
XX 19-JUN-2002; 2002US-00175737.  
XX 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079684P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.

```
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0085023P.
PR 22-MAY-1998; 98US-0085392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089922P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090423P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095988P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096857P.

PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 26-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101933P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Alignment Scores:
Pred. No.: 147
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.63%
DB: 7
Length: 1734
Matches: 8
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-690-825-34 (1-142) x ACD03043 (1-1734)
QY 1 MetGlyAlaProThrLeuProPro 8
DB 1534 ATGGAGCCCCCACACTCCCTCCT 1557

RESULT 68
ACD01858
ID ACD01858 standard; cDNA; 1734 BP.
XX ACD01858;
XX ACD01858;
DT 30-JUL-2003 (first entry)
```



XX Novel human secreted and transmembrane protein PRO1411 cDNA.  
DE Human; secreted and transmembrane protein; PRO; cytostatic; gene therapy;  
XX Chondrocyte stimulator; tumour; adrenal tumour; lung tumour;  
KW colon tumour; breast tumour; prostate tumour; rectal tumour;  
KW cervical tumour; liver tumour; chromosome identification; gene; ss.  
XX Homo sapiens.  
OS  
XX  
XX US2003017544-A1.  
PD  
XX  
XX 23-JAN-2003.  
PF  
XX  
XX 21-JUN-2002; 2002US-00176915.  
PR 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 11-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 12-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 15-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 28-APR-1998; 98US-0083495P.  
PR 28-APR-1998; 98US-0083496P.  
PR 28-APR-1998; 98US-0083499P.  
PR 28-APR-1998; 98US-0083559P.  
PR 03-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 05-JUN-1998; 98US-0088555P.  
PR 05-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 02-JUL-1998; 98US-0091633P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096765P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.



PR	04-JUN-1998	98US-0088025P
PR	04-JUN-1998	98US-0088028P
PR	04-JUN-1998	98US-0088029P
PR	04-JUN-1998	98US-0088033P
PR	04-JUN-1998	98US-0088326P
PR	04-JUN-1998	98US-0088367P
PR	05-JUN-1998	98US-0088402P
PR	05-JUN-1998	98US-0088412P
PR	05-JUN-1998	98US-0088417P
PR	05-JUN-1998	98US-0088455P
PR	10-JUN-1998	98US-0088722P
PR	10-JUN-1998	98US-0088738P
PR	10-JUN-1998	98US-0088740P
PR	10-JUN-1998	98US-0088811P
PR	10-JUN-1998	98US-0088824P
PR	10-JUN-1998	98US-0088825P
PR	10-JUN-1998	98US-0088826P
PR	11-JUN-1998	98US-0088861P
PR	11-JUN-1998	98US-0088863P
PR	11-JUN-1998	98US-0088876P
PR	12-JUN-1998	98US-0089090P
PR	12-JUN-1998	98US-0089105P
PR	16-JUN-1998	98US-0089512P
PR	16-JUN-1998	98US-0089514P
PR	17-JUN-1998	98US-0089538P
PR	17-JUN-1998	98US-0089598P
PR	17-JUN-1998	98US-0089653P
PR	18-JUN-1998	98US-0089908P
PR	19-JUN-1998	98US-0089952P
PR	22-JUN-1998	98US-0090246P
PR	22-JUN-1998	98US-0090252P
PR	22-JUN-1998	98US-0090254P
PR	24-JUN-1998	98US-0090429P
PR	24-JUN-1998	98US-0090435P
PR	24-JUN-1998	98US-0090441P
PR	24-JUN-1998	98US-0090446P
PR	24-JUN-1998	98US-0090535P
PR	24-JUN-1998	98US-0090540P
PR	25-JUN-1998	98US-0090676P
PR	25-JUN-1998	98US-0090678P
PR	25-JUN-1998	98US-0090682P
PR	25-JUN-1998	98US-0090688P
PR	25-JUN-1998	98US-0090694P
PR	25-JUN-1998	98US-0090699P
PR	25-JUN-1998	98US-0090695P
PR	26-JUN-1998	98US-00105413
PR	26-JUN-1998	98US-0090862P
PR	26-JUN-1998	98US-0090863P
PR	01-JUL-1998	98US-0091010P
PR	01-JUL-1998	98US-0091015P
PR	02-JUL-1998	98US-0091478P
PR	02-JUL-1998	98US-0091486P
PR	02-JUL-1998	98US-0091626P
PR	02-JUL-1998	98US-0091628P
PR	02-JUL-1998	98US-0091633P
PR	02-JUL-1998	98US-0094006P
PR	04-AUG-1998	98US-0095282P
PR	10-AUG-1998	98US-0095598P
PR	17-AUG-1998	98US-0096512P
PR	17-AUG-1998	98US-0096757P
PR	17-AUG-1998	98US-0096765P
PR	17-AUG-1998	98US-0096861P
PR	17-AUG-1998	98US-0096867P
PR	17-AUG-1998	98US-0096897P
PR	18-AUG-1998	98US-0096949P
PR	18-AUG-1998	98US-0096959P
PR	18-AUG-1998	98US-0097022P
PR	26-AUG-1998	98US-0097952P
PR	26-AUG-1998	98US-0097954P
PR	26-AUG-1998	98US-0097955P
PR	26-AUG-1998	98US-0097971P
PR	26-AUG-1998	98US-0097974P

PR	26-AUG-1998,	98US-0098014P,
PR	01-SEP-1998,	98US-0098716P,
PR	01-SEP-1998,	98US-0098723P,
PR	02-SEP-1998,	98US-0098803P,
PR	02-SEP-1998,	98US-0098821P,
PR	02-SEP-1998,	98US-0098843P,
PR	03-SEP-1998,	98US-0093602P,
PR	10-SEP-1998,	98US-0093741P,
PR	10-SEP-1998,	98US-0093754P,
PR	10-SEP-1998,	98US-0093763P,
PR	10-SEP-1998,	98US-0093812P,
PR	11-SEP-1998,	98US-0100388P,
PR	16-SEP-1998,	98US-0100662P,
PR	16-SEP-1998,	98US-0100664P,
PR	16-SEP-1998,	98US-0101751P,
PR	16-SEP-1998,	98NO-05019330,
PR	17-SEP-1998,	98US-0101068P,
PR	17-SEP-1998,	98US-0100684P,
PR	17-SEP-1998,	98US-0100931P,
PR	17-SEP-1998,	98US-0100930P,
PR	18-SEP-1998,	98US-0100849P,
PR	18-SEP-1998,	98US-0101014P,
PR	18-SEP-1998,	98US-0101068P,
PR	23-SEP-1998,	98US-0101471P,
PR	23-SEP-1998,	98US-0101472P,
PR	23-SEP-1998,	98US-0101475P,
PR	23-SEP-1998,	98US-0101477P,
PR	24-SEP-1998,	98US-0101738P,
PR	24-SEP-1998,	98US-0101739P,
PR	24-SEP-1998,	98US-0101743P,
PR	24-SEP-1998,	98US-0101822P,
PR	25-SEP-1998,	98US-0101786P,
PR	25-SEP-1998,	98US-0102207P,
PR	25-SEP-1998,	98US-0102240P,
PR	29-SEP-1998,	98US-0102330P,
PR	29-SEP-1998,	98US-0102331P,
PR	30-SEP-1998,	98US-0102487P,
PR	30-SEP-1998,	98US-0102570P,
PR	30-SEP-1998,	98US-0102571P,
PR	01-OCT-1998,	98US-0102684P,
PR	01-OCT-1998,	98US-0102687P,
PR	02-OCT-1998,	98US-0102965P,
PR	06-OCT-1998,	98US-0103258P,
PR	06-OCT-1998,	98US-0103449P,
PR	07-OCT-1998,	98US-0106997P,
PR	07-OCT-1998,	98US-0103395P,
PR	07-OCT-1998,	98US-01033401P,

Alignment Scores:		
Pred. No.:	147	1734
Score:	9.00	8
Percent Similarity:	100.00%	
Best Local Similarity:	100.00%	
Query Match:	5.63%	
Ds:	7	
Length:		
Matches:		0
Conservative:		0
Mismatches:		0
Indels:		0
Gaps:		0

URS-09-690-825-34 (1-142) x ACA92050 (1-1734)

QY  
1 MetGlyAlaProThrLeuProPro 8  
|||  
pb 1534 ATGGAGCCCCCACACTCCCTCCT 1557

## RESULT 70

ACA89475

ID ACA89475 standard; cDNA; 1734 BP.

XX

AC ACA89475;

XX DT 09-JUL-2003 (first entry)

DI XX  
09-00-ET-000 (ATTN: CANTY) /

DE cDNA encoding human PRO polypeptide #101.

100

KW Human; PRO polypeptide; secreted protein; transmembrane protein;

KW chromosome mapping; gene mapping; tumour; adrenal; lung; colon; breast;  
KW prostate; rectal; cervical; liver; cancer; TNF-alpha;  
KW tumour necrosis factor-alpha; proliferation; differentiation;  
KW chondrocyte cell; bone disorder; cartilage disorder; sports injury;  
KW arthritis; cytostatic; antiarthritic; osteopathic; gene therapy; gene  
ss.

XX Homo sapiens.

XX US2003036141-A1.

XX 20-FEB-2003.

XX 01-JUL-2002; 2002US-00187537.

PR 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0082250P.  
PR 21-OCT-1997; 97US-0083486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0083564P.  
PR 29-OCT-1997; 97US-0083734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069333P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-00795664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 15-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.

PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 05-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088867P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090439P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 25-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.

PR	26-AUG-1998;	98US-00979554P.	XX	Homo sapiens.	02-JUL-2002;	2002US-00187603.
PR	26-AUG-1998;	98US-00979555P.	OS		26-JUN-1998;	98US-00105413.
PR	26-AUG-1998;	98US-00979711P.	XX		16-SEP-1998;	98WO-US019330.
PR	26-AUG-1998;	98US-00979749P.	PN	US2003036146-A1.	07-OCT-1998;	98US-00168978.
PR	26-AUG-1998;	98US-00980145P.	XX		07-OCT-1998;	98WO-US021141.
PR	01-SEP-1998;	98US-00987185P.	XX		06-NOV-1998;	98US-00187368.
PR	01-SEP-1998;	98US-00987233P.	PD	20-FEB-2003.	01-DEC-1998;	98WO-US025108.
PR	02-SEP-1998;	98US-00988033P.	XX		01-DEC-1998;	98US-00202054.
PR	02-SEP-1998;	98US-00988212P.	PF		03-MAR-1999;	99US-00254311.
PR	02-SEP-1998;	98US-00988433P.	PR		08-MAR-1999;	99WO-US005028.
PR	09-SEP-1998;	98US-00986032P.	PR		14-MAY-1999;	99US-00311832.
PR	10-SEP-1998;	98US-00987411P.	PR		14-MAY-1999;	99WO-US010733.
PR	10-SEP-1998;	98US-00987633P.	PR		02-JUN-1999;	99WO-US012252.
PR	10-SEP-1998;	98US-00988122P.	PR		25-AUG-1999;	99US-00380137.
PR	15-SEP-1998;	98US-01003889P.	PR		25-AUG-1999;	99US-00380138.
PR	16-SEP-1998;	98US-01005622P.	PR		25-AUG-1999;	99US-00380139.
PR	16-SEP-1998;	98US-01006643P.	PR		25-AUG-1999;	99US-00380142.
PR	16-SEP-1998;	98US-01017541P.	PR		01-SEP-1999;	99WO-US020111.
PR	16-SEP-1998;	98WO-US019330.	PR		15-SEP-1999;	99WO-US021090.
PR	17-SEP-1998;	98US-01006833P.	PR		18-OCT-1999;	99US-00403297.
PR	17-SEP-1998;	98US-01006844P.	PR		12-NOV-1999;	99US-00423844.
PR	17-SEP-1998;	98US-01009199P.	PR		01-DEC-1999;	99WO-US028301.
PR	17-SEP-1998;	98US-01009303P.	PR		02-DEC-1999;	99WO-US028551.
PR	18-SEP-1998;	98US-01008439P.	PR		30-DEC-1999;	99WO-US031274.
PR	18-SEP-1998;	98US-01010147P.	PR		05-JAN-2000;	2000WO-US000219.
PR	18-SEP-1998;	98US-01010688P.	PR		18-FEB-2000;	2000WO-US0004341.
PR	23-SEP-1998;	98US-01014711P.	PR		18-FEB-2000;	2000WO-US0004342.
PR	23-SEP-1998;	98US-01014722P.	PR		22-FEB-2000;	2000WO-US0004414.
PR	23-SEP-1998;	98US-01014759P.	PR		24-FEB-2000;	2000WO-US005004.
PR	23-SEP-1998;	98US-01014779P.	PR		01-MAR-2000;	2000WO-US005601.
PR	24-SEP-1998;	98US-01017388P.	PR		02-MAR-2000;	2000WO-US005841.
PR	24-SEP-1998;	98US-01017399P.	PR		15-MAR-2000;	2000WO-US006884.
PR	24-SEP-1998;	98US-01017433P.	PR		30-MAR-2000;	2000WO-US008439.
PR	24-SEP-1998;	98US-01017434P.	PR		17-MAY-2000;	2000WO-US013705.
PR	25-SEP-1998;	98US-01019222P.	PR		22-MAY-2000;	2000WO-US014042.
PR	25-SEP-1998;	98US-01017869P.	PR		30-MAY-2000;	2000WO-US014941.
PR	29-SEP-1998;	98US-01023076P.	PR		02-JUN-2000;	2000WO-US015264.
PR	29-SEP-1998;	98US-01023207P.	PR		28-JUL-2000;	2000WO-US020710.
PR	29-SEP-1998;	98US-01023403P.	PR		22-AUG-2000;	2000US-00644848.
PR	29-SEP-1998;	98US-01023331P.	PR		24-AUG-2000;	2000WO-US023328.
PR	30-SEP-1998;	98US-01024879P.	PR		18-SEP-2000;	2000US-00664610.
PR	30-SEP-1998;	98US-01025707P.	PR		18-SEP-2000;	2000US-00665350.
PR	30-SEP-1998;	98US-01025711P.	PR		08-NOV-2000;	2000US-00709238.
PR	01-OCT-1998;	98US-01026584P.	PR		08-NOV-2000;	2000WO-US030952.
PR	01-OCT-1998;	98US-01026879P.	PR		01-DEC-2000;	2000WO-US032678.
PR	01-OCT-1998;	98US-01026879P.	PR		20-DEC-2000;	2000US-00747259.
PR	01-OCT-1998;	98US-01026879P.	PR		20-DEC-2000;	2000WO-US034956.
PR	01-OCT-1998;	98US-01026879P.	PR		28-DEC-2000;	2000WO-US006520.
PR	01-OCT-1998;	98US-01026879P.	PR		22-MAR-2001;	2001US-00816744.
PR	01-OCT-1998;	98US-01026879P.	PR		10-MAY-2001;	2001US-00854208.
PR	01-OCT-1998;	98US-01026879P.	PR		25-MAY-2001;	2001US-00854280.
PR	01-OCT-1998;	98US-01026879P.	PR		01-JUN-2001;	2001US-00866028.
PR	01-OCT-1998;	98US-01026879P.	PR		05-JUN-2001;	2001WO-US017800.
PR	01-OCT-1998;	98US-01026879P.	PR		20-JUN-2001;	2001US-00874503.
PR	01-OCT-1998;	98US-01026879P.	PR		20-JUN-2001;	2001WO-US019692.
PR	01-OCT-1998;	98US-01026879P.	PR		29-JUN-2001;	2001WO-US021066.
PR	01-OCT-1998;	98US-01026879P.	PR		09-JUL-2001;	2001WO-US021735.
PR	01-OCT-1998;	98US-01026879P.	PR		18-JUL-2001;	2001US-00908827.
PR	01-OCT-1998;	98US-01026879P.	PR		30-JUL-2001;	2001US-00918585.
PR	01-OCT-1998;	98US-01026879P.	PR		06-AUG-2001;	2001US-00924419.
PR	01-OCT-1998;	98US-01026879P.	PR		13-AUG-2001;	2001US-00929404.
PR	01-OCT-1998;	98US-01026879P.	PR		16-AUG-2001;	2001US-00931836.
PR	01-OCT-1998;	98US-01026879P.	PR		28-AUG-2001;	2001US-00941992.
PR	01-OCT-1998;	98US-01026879P.	PR		29-AUG-2001;	2001WO-US027099.
PR	01-OCT-1998;	98US-01026879P.	PR		04-SEP-2001;	2001US-00946374.

Alignment Scores: 147 1734  
Pred. No.: 8.00  
Score: 100.00%  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 5.63%  
DB: 7

US-09-690-825-34 (1-142) x ACA89475 (1-1734)

QY 1 MetGlyAlaProThrLeuPro 8  
DB 1534 ATCGGAGCCCCACACTCCCTCT 1557

RESULT 71

ACA73485

ID ACA73485 standard; cDNA; 1734 BP.

XX ACA73485;

AC ACA73485;

XX 01-JUL-2003 (first entry)

DT Human secreted/transmembrane protein (PRO) cDNA #101.

DE Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;  
XX proliferation; differentiation; chondrocyte cells;  
KW tumour necrosis factor-alpha; TNF-alpha; blood; gene therapy.

PR 15-JAN-2002; 2002US-00052586.  
 PA (GETH ) GENENTECH INC.  
 XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;  
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;  
 XX WPI; 2003-332034/31.  
 DR P-PSDB; AB086293.  
 XX  
 XX Three hundred and five nucleic acids encoding PRO polypeptides, useful in  
 PT gene therapy, chromosome identification, tissue typing, and for detecting  
 PT the presence of tumor in a mammal.  
 XX  
 XX Claim 2; Fig 201; 707pp; English.  
 XX  
 CC The invention relates to three hundred and five nucleic acids encoding  
 CC PRO polypeptides (secreted and transmembrane), sequences 80% identical to  
 CC them, or encoding a PRO polypeptide lacking its associated signal peptide  
 CC or an extracellular domain of the PRO polypeptide, with or lacking its  
 CC associated signal peptide. Also included are the encoded PRO proteins,  
 CC PRO expression vectors, host cells transformed with the vector (used to  
 CC produce PRO proteins), a chimeric molecule comprising the PRO  
 CC polypeptide fused to a heterologous amino acid sequence, an anti-PRO  
 CC antibody, a method for stimulating the release of tumor necrosis factor  
 CC alpha (TNF-alpha) from human blood (by contacting the blood with PRO1079,  
 CC PRO827, PRO791, PRO1131, PRO1316, PRO1183, PRO1343, PRO1760, PRO1567 or  
 CC PRO4333), a method for stimulating the proliferation or differentiation  
 CC of chondrocyte cells by contacting the cells with a PRO6029 polypeptide,  
 CC a method for detecting the presence of tumor in a mammal and an  
 CC oligonucleotide probe derived from any of the nucleotide sequences cited  
 CC above. The PRO polypeptide or anti-PRO antibody is useful for preparing a  
 CC medicament for treating a condition that is responsive to the PRO  
 CC polypeptide or anti-PRO antibody. The PRO nucleotide sequences are useful  
 CC as hybridisation probes in chromosome and gene mapping, or in generating  
 CC antisense RNA and DNA. PRO nucleic acids are also useful in preparing PRO  
 CC polypeptides, in assays to identify other proteins or molecules involved  
 CC in a binding reaction, to generate transgenic animals or knockout  
 CC animals, which in turn are useful in the development and screening of  
 CC therapeutically useful reagents, for chromosome identification, and  
 CC tissue typing. The PRO polypeptides and nucleic acid molecules are also  
 CC useful for detecting the presence of a tumour in a mammal, stimulating the  
 CC proliferation or differentiation of chondrocyte cells, stimulating the  
 CC release of tumour necrosis factor-alpha from human blood, in gene  
 CC therapy, or as molecular weight markers for protein electrophoresis  
 CC purposes. The anti-PRO antibodies may be used in diagnostic assays for  
 CC PRO, or for the affinity purification of PRO from recombinant cell  
 CC culture or natural sources. The present sequence is a cDNA encoding a PRO  
 CC protein  
 XX  
 SQ Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 147 Length: 1734  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.63% Indels: 0  
 DB: 7 Gaps: 0

US-09-690-825-34 (1-142) x ACA73485 (1-1734)

QV 1 MetGlyAlaProThrLeuProPro 8  
 Db 1534 ATGGAGCCCCACACTCCCTCT 1557

RESULT 72

ACA05800

ID ACA05800 standard; cDNA; 1734 BP.

XX

AC ACA05800;

XX

DT 29-MAY-2003 (first entry)

XX Human secreted/transmembrane protein (PRO) cDNA #101.  
 DE  
 XX Human; gene; ss; secreted and transmembrane protein; PRO; TNF-alpha;  
 KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;  
 KW tissue typing.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003036162-A1.  
 XX  
 PD 20-FEB-2003.  
 XX  
 PF 12-JUL-2002; 2002US-00194423.  
 XX  
 PR 26-JUN-1998; 98US-00105413.  
 PR 16-SEP-1998; 98WO-US019330.  
 PR 07-OCT-1998; 98US-00168978.  
 PR 07-OCT-1998; 98WO-US021141.  
 PR 06-NOV-1998; 98US-00187368.  
 PR 01-DEC-1998; 98WO-US025108.  
 PR 07-DEC-1998; 98US-00202054.  
 PR 03-MAR-1999; 98US-00254311.  
 PR 08-MAR-1999; 98WO-US005028.  
 PR 14-MAY-1999; 98US-00311832.  
 PR 14-MAY-1999; 98WO-US010733.  
 PR 02-JUN-1999; 98US-0012252.  
 PR 25-AUG-1999; 98US-00380137.  
 PR 25-AUG-1999; 98US-00380138.  
 PR 25-AUG-1999; 98US-00380139.  
 PR 25-AUG-1999; 98US-00380142.  
 PR 01-SEP-1999; 98WO-US020111.  
 PR 15-SEP-1999; 98WO-US021090.  
 PR 18-OCT-1999; 98US-00403297.  
 PR 12-NOV-1999; 98US-00423844.  
 PR 01-DEC-1999; 98WO-US028301.  
 PR 02-DEC-1999; 98WO-US028551.  
 PR 30-DEC-1999; 98WO-US031274.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 22-FEB-2000; 2000WO-US004342.  
 PR 24-FEB-2000; 2000WO-US004414.  
 PR 01-MAR-2000; 2000WO-US005004.  
 PR 02-MAR-2000; 2000WO-US005601.  
 PR 15-MAR-2000; 2000WO-US005841.  
 PR 30-MAR-2000; 2000WO-US006884.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 28-JUL-2000; 2000WO-US020710.  
 PR 22-AUG-2000; 2000US-00844848.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 18-SEP-2000; 2000US-00664610.  
 PR 18-SEP-2000; 2000US-00665350.  
 PR 08-NOV-2000; 2000US-00709238.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 20-DEC-2000; 2000US-00747259.  
 PR 20-DEC-2000; 2000WO-US034956.  
 PR 28-FEB-2001; 2001WO-US006520.  
 PR 22-MAR-2001; 2001US-00816744.  
 PR 10-MAY-2001; 2001US-00854208.  
 PR 10-MAY-2001; 2001US-00854280.  
 PR 25-MAY-2001; 2001US-00866028.  
 PR 01-JUN-2001; 2001WO-US017800.  
 PR 05-JUN-2001; 2001US-00874503.  
 PR 20-JUN-2001; 2001WO-US019692.  
 PR 29-JUN-2001; 2001WO-US021066.  
 PR 09-JUL-2001; 2001WO-US021735.  
 PR 18-JUL-2001; 2001US-00908827.  
 PR 30-JUL-2001; 2001US-00918585.

PR 06-AUG-2001; 2001US-00924419.  
 PR 13-AUG-2001; 2001US-00929404.  
 PR 16-AUG-2001; 2001US-00931836.  
 PR 28-AUG-2001; 2001US-00941992.  
 PR 29-AUG-2001; 2001WO-US027099.  
 PR 04-SEP-2001; 2001US-00946374.  
 PR 15-JAN-2002; 2002US-00052586.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;  
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;  
 XX  
 XX WPI; 2003-332039/31.  
 DR P-PSDB; ABU67506.  
 XX  
 XX  
 PT New secreted and transmembrane PRO polypeptides and nucleic acids, useful  
 PT in gene therapy, in chromosome and gene mapping, as chromosome markers,  
 PT in tissue typing, and in chromosome identification.  
 XX  
 PS Claim 2; Fig 201; 706pp; English.  
 XX  
 CC The invention discloses human nucleic acids encoding secreted and  
 CC transmembrane (PRO) polypeptides. Also disclosed is an antibody that  
 CC specifically binds to the PRO polypeptide, a method for stimulating the  
 CC release of tumour necrosis factor alpha (TNF-alpha) from human blood by  
 CC contacting the blood a PRO polypeptide, a method for stimulating the  
 CC proliferation or differentiation of chondrocyte cells by contacting the  
 CC cells with a PRO polypeptide, a method for detecting the presence of a  
 CC tumour in a mammal and an oligonucleotide probe derived from any of the  
 CC PRO nucleotide sequences. The nucleotide sequences are useful as probes,  
 CC in chromosome and gene mapping, in generating antisense RNA and DNA, in  
 CC preparing PRO polypeptides by recombinant techniques and in gene therapy  
 CC (e.g. for replacement of defective gene). The PRO polypeptides are useful  
 CC as molecular weight markers for protein electrophoresis purposes, for  
 CC chromosome identification, as chromosome markers, as therapeutic agents,  
 CC for stimulating the release of TNF-alpha from human blood, for  
 CC stimulating the proliferation or differentiation of chondrocytes and  
 CC detecting the presence of a tumour. The PRO polypeptides and nucleic  
 CC acids may also be used diagnostically for tissue typing. The sequences  
 CC presented in ACA05700-ACA06004 are the cDNAs encoding the PRO  
 CC polypeptides of the invention  
 XX  
 SQ Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 147 Length: 1734  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.63% Indels: 0  
 DB: 7 Gaps: 0  
 US-09-690-825-34 (1-142) x ACA05800 (1-1734)  
 QY 1 MetGlyAlaProthrLeuproPro 8  
 DB 1534 ATGGAGGCCCCACATCCCTCTCT 1557  
 RESULT 73  
 ID ACA66634  
 XX ID ACA66634 standard; cDNA; 1734 BP.  
 AC ACA66634;  
 XX  
 DT 23-JUN-2003 (first entry)  
 XX cDNA encoding human PRO protein #101.  
 XX Human; tumour; adrenal; lung; colon; breast; prostate; rectal; cervical;  
 KW liver; PRO; gene therapy; gene; ss.  
 XX  
 OS Homo sapiens.

XX US2003036137-A1.  
 XX  
 PD 20-FEB-2003.  
 XX  
 PF 27-JUN-2002; 2002US-00184640.  
 XX  
 PR 26-JUN-1998; 98US-00105413.  
 PR 16-SEP-1998; 98WO-US019330.  
 PR 07-OCT-1998; 98US-00168978.  
 PR 07-OCT-1998; 98WO-US021141.  
 PR 06-NOV-1998; 98US-00187368.  
 PR 01-DEC-1998; 98WO-US025108.  
 PR 07-DEC-1998; 98US-00202054.  
 PR 03-MAR-1999; 99US-00254311.  
 PR 08-MAR-1999; 99WO-US005028.  
 PR 14-MAY-1999; 99US-00311832.  
 PR 14-MAY-1999; 99WO-US010733.  
 PR 02-JUN-1999; 99WO-US012252.  
 PR 25-AUG-1999; 99US-00380137.  
 PR 25-AUG-1999; 99US-00380138.  
 PR 25-AUG-1999; 99US-00380139.  
 PR 25-AUG-1999; 99US-00380142.  
 PR 01-SEP-1999; 99WO-US020111.  
 PR 15-SEP-1999; 99WO-US021090.  
 PR 18-OCT-1999; 99US-00403297.  
 PR 12-NOV-1999; 99US-00423844.  
 PR 01-DEC-1999; 99WO-US028301.  
 PR 30-DEC-1999; 99WO-US028551.  
 PR 30-DEC-1999; 99WO-US031274.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 18-FEB-2000; 2000WO-US004342.  
 PR 22-FEB-2000; 2000WO-US004414.  
 PR 24-FEB-2000; 2000WO-US005004.  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 28-JUL-2000; 2000WO-US020710.  
 PR 22-AUG-2000; 2000WO-US024848.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 18-SEP-2000; 2000US-00664610.  
 PR 18-SEP-2000; 2000US-00665350.  
 PR 08-NOV-2000; 2000US-00709238.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 20-DEC-2000; 2000US-00747259.  
 PR 20-DEC-2000; 2000WO-US034956.  
 PR 28-FEB-2001; 2001WO-US006520.  
 PR 22-MAR-2001; 2001US-00816744.  
 PR 10-MAY-2001; 2001US-00854208.  
 PR 10-MAY-2001; 2001US-00854280.  
 PR 25-MAY-2001; 2001US-00866028.  
 PR 01-JUN-2001; 2001WO-US017800.  
 PR 05-JUN-2001; 2001US-00874503.  
 PR 20-JUN-2001; 2001WO-US019692.  
 PR 29-JUN-2001; 2001WO-US021066.  
 PR 09-JUL-2001; 2001WO-US021735.  
 PR 30-JUL-2001; 2001US-00908827.  
 PR 30-JUL-2001; 2001US-00918595.  
 PR 06-AUG-2001; 2001US-00924419.  
 PR 13-AUG-2001; 2001US-00929404.  
 PR 16-AUG-2001; 2001US-00931836.  
 PR 28-AUG-2001; 2001US-00941992.  
 PR 29-AUG-2001; 2001WO-US027099.  
 PR 04-SEP-2001; 2001US-00946374.  
 PR 15-JAN-2002; 2002US-00052586.  
 XX

PA (GETH ) GENENTECH INC.  
 XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;  
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;  
 XX WPI; 2003-342038/32.  
 DR P-PSDB; ABU0534.  
 XX Three hundred and five nucleic acids encoding secreted and transmembrane  
 PT PRO polypeptides, useful for the diagnosis, prevention and/or treatment  
 PT of tumors, such as adrenal, lung, colon, breast, prostate, rectal,  
 PT cervical or liver tumors.  
 XX Claim 2; Fig 201; 708pp; English.  
 PS The invention relates to three hundred and five nucleic acids encoding  
 CC PRO polypeptides (secreted and transmembrane). Methods and compositions  
 CC of the present invention are useful for the diagnosis, prevention and/or  
 CC treatment of tumors, such as adrenal, lung, colon, breast, prostate,  
 CC rectal, cervical or liver tumors. The PRO polypeptides are also useful  
 CC as molecular weight markers, or for chromosome identification. The PRO  
 CC genes are useful as hybridisation probes, or for screening libraries of  
 CC human cDNA, genomic DNA or mRNA. The PRO genes may also be used in gene  
 CC therapy, particularly for replacing a defective gene. The present  
 CC sequence represents a cDNA encoding a human PRO polypeptide of the  
 CC invention  
 XX Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;  
 SQ

Alignment Scores:  
 Pred. No.: 147 Length: 1734  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.63% Indels: 0  
 DB: 7 Gaps: 0

US-09-690-825-34 (1-142) x ACA66534 (1-1734)  
 QY 1 MetGlyAlaProThrLeuProPro 8  
 Db 1534 ATGGAGGCCCCACACTCCCTCCT 1557

RESULT 74  
 AC91189  
 ID AC91189 standard; cDNA; 1734 BP.  
 AC ACA91189;  
 XX 11-JUL-2003 (first entry)  
 DT Novel human secreted and transmembrane protein PRO1411 cDNA.  
 DE Human; secreted and transmembrane protein; PRO; antibody therapy;  
 KW pharmaceutical; diagnostic; biosensor; bioreactor; gene; ss.  
 XX Homo sapiens.  
 OS US2003018173-A1.  
 XX 23-JAN-2003.  
 PD 01-MAY-2002; 2002US-00063515.  
 PF 06-DEC-2001; 2001US-00006867.  
 PR (GETH ) GENENTECH INC.  
 PA Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;  
 PI Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;  
 XX WPI; 2003-401702/38.  
 DR P-PSDB; ABU0901.  
 XX New antibody useful for identifying PRO polypeptides, for affinity  
 PT purification of PRO polypeptides, and for preparing a medicament for  
 PT diagnosing or treating conditions responsive to the antibody or PRO  
 PT polypeptide.  
 XX Disclosure; Fig 51; 345pp; English.  
 PS The invention describes an antibody that specifically binds to a PRO  
 CC polypeptide having a fully defined amino acid sequence given in the  
 CC specification. The antibody is useful in identifying PRO polypeptides  
 CC useful for various industrial applications, including pharmaceuticals,  
 CC diagnostics, biosensors and bioreactors. The antibody is also used for  
 CC affinity purification of PRO polypeptides from recombinant cell culture  
 CC or natural sources. The antibody, PRO polypeptide, or its agonists or  
 CC antagonists, may be used for preparing a medicament for diagnosing or  
 CC treating a condition responsive to the antibody, PRO polypeptide, or its  
 CC agonists or antagonists. This sequence encodes a novel human secreted and  
 CC transmembrane PRO polypeptide  
 XX Sequence 1734 BP; 503 A; 397 C; 580 G; 254 T; 0 U; 0 Other;  
 SQ

Alignment Scores:  
 Pred. No.: 147 Length: 1734  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.63% Indels: 0  
 DB: 7 Gaps: 0

US-09-690-825-34 (1-142) x ACA91189 (1-1734)  
 QY 1 MetGlyAlaProThrLeuProPro 8  
 Db 1534 ATGGAGGCCCCACACTCCCTCCT 1557

RESULT 75  
 AC91189  
 ID AC91189 standard; cDNA; 1734 BP.  
 AC ACD81566;  
 XX 18-SEP-2003 (first entry)  
 DT Human cDNA encoding secreted/transmembrane protein PRO411.  
 DE Human; ss; gene; secreted/transmembrane protein; PRO; tumour; cancer;  
 KW cytostatic.  
 XX Homo sapiens.  
 OS US2003009013-A1.  
 XX 09-JAN-2003.  
 PD 01-MAY-2002; 2002US-00063519.  
 PF 30-DEC-1998; 98KR-00062142.  
 PR 08-MAR-1999; 99WO-US005028.  
 PR 14-MAY-1999; 99US-00311832.  
 PR 14-MAY-1999; 99WO-US010733.  
 PR 25-AUG-1999; 99US-00380137.  
 PR 25-AUG-1999; 99US-00380138.  
 PR 25-AUG-1999; 99US-00380139.  
 PR 25-AUG-1999; 99US-00380142.  
 PR 15-SEP-1999; 99US-00397342.  
 PR 18-OCT-1999; 99US-00403297.  
 PR 12-NOV-1999; 99US-00423844.  
 PR 30-DEC-1999; 99WO-US031274.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 21-MAR-2000; 2000WO-US007532.



Search completed: August 16, 2004, 01:35:39  
Job time : 483 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 16, 2004, 02:47:05 ; Search time 1465 Seconds

(without alignments)  
475.589 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 142  
Sequence: 1 MGATLPAPQWFLKDHRS.....EFETAKKVRRAIEQLAAMD 142

Scoring table:

Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6442761

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODE=frame+pn2.model -DRV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09690825/runat\_11082004\_141017\_13907/app\_query.fasta\_1.327  
-DB=Published Applications NA -QWFASTAP -SUFFIX=oligo.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=100 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1  
-ALIGN=75 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09690825 -CGN1\_1\_520 @runat\_11082004\_141017\_13907  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description

1	142	100.0	429	17	US-10-665-975-1	Sequence 1, Appl
2	142	100.0	1619	9	US-09-918-186A-97	Sequence 97, Appl
3	142	100.0	1619	13	US-10-342-887-566	Sequence 566, App
4	142	100.0	1619	13	US-10-172-118-566	Sequence 566, App
5	142	100.0	1619	13	US-10-181-316-97	Sequence 97, Appl
6	142	100.0	1619	13	US-10-388-360-324	Sequence 324, App
7	142	100.0	1619	17	US-10-283-975A-445	Sequence 445, App
8	142	100.0	2404	14	US-10-071-766-109	Sequence 109, App
9	142	100.0	2404	15	US-10-084-817-285	Sequence 285, App
10	93	65.5	740	16	US-10-305-720-918	Sequence 918, App
11	41	28.9	955	9	US-09-918-186A-10	Sequence 10, Appl
12	41	28.9	955	13	US-10-181-316-10	Sequence 10, Appl
13	40	28.2	121	15	US-10-179-730-33	Sequence 33, Appl
14	39	27.5	14795	13	US-10-181-316-3	Sequence 3, Appl
15	39	27.5	14796	9	US-09-954-456-973	Sequence 973, Appl
16	39	27.5	14796	9	US-09-954-456-1636	Sequence 1636, Ap
17	39	27.5	14796	9	US-09-918-186A-3	Sequence 3, Appl
18	39	27.5	14796	9	US-09-880-107-3421	Sequence 3421, Ap
19	39	27.5	14796	12	US-09-968-007A-522	Sequence 522, App
20	39	27.5	14796	15	US-10-138-618-35	Sequence 35, Appl
21	36	25.4	120	15	US-10-029-386-23672	Sequence 23672, A
22	36	25.4	555	15	US-10-029-386-9972	Sequence 9972, Ap
23	36	25.4	1100	15	US-10-108-877-2	Sequence 2, Appl
24	14	9.9	332	9	US-09-833-381-1331	Sequence 1331, Ap
25	9	6.3	30	15	US-10-400-382-265	Sequence 265, App
26	9	6.3	30	17	US-10-636-065-199	Sequence 199, App
27	8	5.6	26	9	US-09-918-186A-13	Sequence 13, Appl
28	8	5.6	26	13	US-10-181-316-13	Sequence 13, Appl
29	8	5.6	281	9	US-09-833-381-1330	Sequence 1330, Ap
30	8	5.6	420	13	US-10-027-632-65868	Sequence 65868, A
31	8	5.6	420	13	US-10-027-632-65870	Sequence 65870, A
32	8	5.6	420	13	US-10-027-632-65870	Sequence 65870, A
33	8	5.6	420	16	US-10-027-632-65868	Sequence 65868, A
34	8	5.6	420	16	US-10-027-632-65869	Sequence 65869, A
35	8	5.6	420	16	US-10-027-632-65870	Sequence 65870, A
36	8	5.6	454	13	US-10-424-599-114486	Sequence 114486, A
37	8	5.6	515	13	US-09-918-186A-231	Sequence 231, App
38	8	5.6	515	13	US-10-181-316-231	Sequence 231, App
39	8	5.6	587	10	US-09-918-995-15078	Sequence 15078, A
40	8	5.6	1224	16	US-10-369-493-35886	Sequence 35886, A
41	8	5.6	1441	15	US-10-050-704-14	Sequence 14, Appl
42	8	5.6	1441	17	US-10-798-512-14	Sequence 14, Appl
43	8	5.6	1734	10	US-09-946-374-51	Sequence 51, Appl
44	8	5.6	1734	12	US-10-015-395A-51	Sequence 51, Appl
45	8	5.6	1734	13	US-10-206-915-201	Sequence 201, App
46	8	5.6	1734	13	US-10-199-670-201	Sequence 201, App
47	8	5.6	1734	13	US-10-201-858-201	Sequence 201, App
48	8	5.6	1734	13	US-10-081-056-149	Sequence 149, App
49	8	5.6	1734	13	US-10-219-535-155	Sequence 155, App
50	8	5.6	1734	13	US-10-232-230-155	Sequence 155, App
51	8	5.6	1734	13	US-10-205-890-201	Sequence 201, App
52	8	5.6	1734	13	US-10-208-024-201	Sequence 201, App
53	8	5.6	1734	13	US-10-201-853-201	Sequence 201, App
54	8	5.6	1734	13	US-10-063-745-51	Sequence 51, Appl
55	8	5.6	1734	13	US-10-063-512-51	Sequence 51, Appl
56	8	5.6	1734	13	US-10-063-515-51	Sequence 51, Appl
57	8	5.6	1734	13	US-10-063-515-51	Sequence 51, Appl
58	8	5.6	1734	13	US-10-063-549-51	Sequence 51, Appl
59	8	5.6	1734	13	US-10-063-569-51	Sequence 51, Appl
60	8	5.6	1734	13	US-10-063-551-51	Sequence 51, Appl
61	8	5.6	1734	13	US-10-174-581-201	Sequence 201, App
62	8	5.6	1734	13	US-10-176-483-201	Sequence 201, App
63	8	5.6	1734	13	US-10-176-749-201	Sequence 201, App
64	8	5.6	1734	13	US-10-176-914-201	Sequence 201, App
65	8	5.6	1734	13	US-10-176-915-201	Sequence 201, App
66	8	5.6	1734	13	US-10-006-485A-51	Sequence 51, Appl
67	8	5.6	1734	13	US-10-013-907A-51	Sequence 51, Appl
68	8	5.6	1734	13	US-10-015-499A-51	Sequence 51, Appl
69	8	5.6	1734	13	US-10-063-555-51	Sequence 51, Appl
70	8	5.6	1734	13	US-10-063-563-51	Sequence 51, Appl
71	8	5.6	1734	13	US-10-063-594-51	Sequence 51, Appl
72	8	5.6	1734	13	US-10-063-553-51	Sequence 51, Appl

73 8 5.6 1734 13 US-10-063-554-51 Sequence 51, Appl  
 74 8 5.6 1734 13 US-10-176-484-201 Sequence 201, App  
 75 8 5.6 1734 13 US-10-180-550-201 Sequence 201, App  
 76 8 5.6 1734 13 US-10-183-014-201 Sequence 201, App  
 77 8 5.6 1734 13 US-10-187-738-201 Sequence 201, App  
 78 8 5.6 1734 13 US-10-187-740-201 Sequence 201, App  
 79 8 5.6 1734 13 US-10-187-883-201 Sequence 201, App  
 80 8 5.6 1734 13 US-10-194-363-201 Sequence 201, App  
 81 8 5.6 1734 13 US-10-194-460-201 Sequence 201, App  
 82 8 5.6 1734 13 US-10-194-463-201 Sequence 201, App  
 83 8 5.6 1734 13 US-10-194-484-201 Sequence 201, App  
 84 8 5.6 1734 13 US-10-195-884-201 Sequence 201, App  
 85 8 5.6 1734 13 US-10-195-896-201 Sequence 201, App  
 86 8 5.6 1734 13 US-10-196-744-201 Sequence 201, App  
 87 8 5.6 1734 13 US-10-196-755-201 Sequence 201, App  
 88 8 5.6 1734 13 US-10-196-757-201 Sequence 201, App  
 89 8 5.6 1734 13 US-10-197-704-201 Sequence 201, App  
 90 8 5.6 1734 13 US-10-197-710-201 Sequence 201, App  
 91 8 5.6 1734 13 US-10-198-758-201 Sequence 201, App  
 92 8 5.6 1734 13 US-10-198-766-201 Sequence 201, App  
 93 8 5.6 1734 13 US-10-199-304-201 Sequence 201, App  
 94 8 5.6 1734 13 US-10-199-309-201 Sequence 201, App  
 95 8 5.6 1734 13 US-10-199-313-201 Sequence 201, App  
 96 8 5.6 1734 13 US-10-199-456-201 Sequence 201, App  
 97 8 5.6 1734 13 US-10-201-329-201 Sequence 201, App  
 98 8 5.6 1734 13 US-10-202-412-201 Sequence 201, App  
 99 8 5.6 1734 13 US-10-206-919-201 Sequence 201, App  
 100 8 5.6 1734 13 US-10-206-922-201 Sequence 201, App

## ALIGNMENTS

RESULT 1  
 US-10-665-975-1  
 ; Sequence 1, Application US/10665975  
 ; Publication No. US20040138119A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, John C.  
 ; APPLICANT: Tamm, Ingo  
 ; TITLE OF INVENTION: USE OF HEPATITIS VIRUS B X-INTERACTING  
 ; TITLE OF INVENTION: PROTEIN (HEXIP) IN MODULATION OF APOPTOSIS  
 ; FILE REFERENCE: BURNHAM.005A  
 ; CURRENT APPLICATION NUMBER: US/10/665.975  
 ; CURRENT FILING DATE: 2003-09-18  
 ; PRIOR APPLICATION NUMBER: 60/412,109  
 ; PRIOR FILING DATE: 2002-09-18  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 429  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-665-975-1

Alignment Scores:  
 Pred. No.: 7.45e-146 Length: 429  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 17 Gaps: 0

US-09-690-825-34 (1-142) x US-10-665-975-1 (1-429)  
 QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 1 ATGGGTGCCCCGACGTTGCCCTGCGCCCTGGCAGCCCTTCTCAAGGACCACCGCATCTCT 60  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 61 ACATTCAGAACTGGCCCTTCTTGAGGGCTGCGCCCTGCACCCCGGAGCGGATGGCGGAG 120  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60

Db 121 GCTGGCTTCATCCACTGCCCTGAGAACGAGCAGACTTGGCCAGTGTTCTTCTGCG 180  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
 Db 181 TTCAGGAGCTGAAGGCTGGGAGCCAGATGACGACCCCATAGAGGAACATAAAAGCAT 240  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
 Db 241 TCGTCGGTGGCGCTTCCTTCTCTCAAGAAAGCAGTTCGAAGAATTAAACCTTGGTGAA 300  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 Db 301 TTTTGAACCTGACAGAGAGAGCCCAAGAACAAATTTGCAAGGAACCAACAATAAG 360  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
 Db 361 AAGAAAGAAATTTGAGGAAACTCGAAGAAAGTGGCCGTCATCGAGCAGCTGCTGCC 420  
 QY 141 MetAsp 142  
 Db 421 ATGGAT 426

## RESULT 2

US-09-918-186A-97  
 ; Sequence 97, Application US/09918186A  
 ; Patent No. US20020137708A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: C. Frank Bennett  
 ; APPLICANT: Elizabeth J. Ackermann  
 ; APPLICANT: Eric E. Swayze  
 ; APPLICANT: Lex M. Cowsett  
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
 ; FILE REFERENCE: ISPH-0585  
 ; CURRENT APPLICATION NUMBER: US/09/918,186A  
 ; CURRENT FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 09/496,694  
 ; PRIOR FILING DATE: 2000-02-02  
 ; PRIOR APPLICATION NUMBER: 09/286,407  
 ; PRIOR FILING DATE: 1999-04-05  
 ; PRIOR APPLICATION NUMBER: 09/163,162  
 ; PRIOR FILING DATE: 1998-09-29  
 ; NUMBER OF SEQ ID NOS: 250  
 ; SEQ ID NO 97  
 ; LENGTH: 1619  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-918-186A-97

Alignment Scores:  
 Pred. No.: 2.54e-145 Length: 1619  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x US-09-918-186A-97 (1-1619)  
 QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 50 ATGGGTGCCCCGACGTTGCCCTGCGCCCTTCTCAAGGACCACCGCATCTCT 109  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 110 ACATTCAGAACTGGCCCTTCTTGAGGGCTGCGCCCTGCACCCCGGAGCGGATGGCGGAG 169  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
 Db 170 GCTGGCTTCATCCACTGCCCTGAGAACGAGCAGACTTGGCCAGTGTTCTTCTGCG 229  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80

Db 230 TTCAAGGAGCTGGAAGGCTGGGAGCGCAGATGACGACCCCATAGAGAAACATAAAAGCAT 289

Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100

Db 290 TCGTCCGGTTCGCTTTCTTCTGTCAAGAGCAGTTTGAGAAATTAAACCTTGGTGA 349

Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120

Db 350 TTTTGAAGAACTGGACAGAGAAAGCCCAAGAACAAAATTGCAAGAGAAACCAACAATAAG 409

Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluLeuAlaAla 140

Db 410 AAGAAAGAAATTTGAGGAAACTGCGAAGAAAGTGGCGCGTCCCATCGACAGCTGGCTGCC 469

Qy 141 MetAsp 142

Db 470 ATGGAT 475

RESULT 3

US-10-342-887-566

Sequence 566, Application US/10342887

Publication No. US20040058340A1

GENERAL INFORMATION:

APPLICANT: Dai, Hongyue

APPLICANT: He, Yudong

APPLICANT: Linsley, Peter S.

APPLICANT: Mao, Mao

APPLICANT: Roberts, Christopher J.

APPLICANT: Van 't Veer, Laura Johanna

APPLICANT: Bernards, Marc J.

TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

FILE REFERENCE: 9301-188-999

CURRENT APPLICATION NUMBER: US/10/342,887

CURRENT FILING DATE: 2003-01-15

PRIOR APPLICATION NUMBER: 60/298,918

PRIOR FILING DATE: 2001-06-18

PRIOR APPLICATION NUMBER: 60/380,710

PRIOR FILING DATE: 2002-05-14

PRIOR APPLICATION NUMBER: 10/172,118

PRIOR FILING DATE: 2002-06-14

NUMBER OF SEQ ID NOS: 2699

SEQ ID NO 566

LENGTH: 1619

TYPE: DNA

ORGANISM: Homo sapiens

PUBLICATION INFORMATION:

DATABASE ACCESSION NUMBER: NM\_001168

DATABASE ENTRY DATE: 2001-06-18

US-10-342-887-566

Alignment Scores:

Pred. No.: 2,54e-145 Length: 1619

Score: 142.00 Matches: 142

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-342-887-566 (1-1619)

Qy 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20

Db 50 ATGGGTGCCCGCAGCTTGCCTGCGCCCTTCTCAAGGACCAACCGCATCTCT 109

Qy 21 ThrPheLysAsnTrpPropheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40

Db 110 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGGCGCTGCACCCCGGAGCGGATGGCCGAG 169

Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60

Db 170 GCTGGCTTCATCCACTGCCCTTCTTGGAGGGCTGGCGCTGCACCCCGGAGCGGATGGCCGAG 229

Qy 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80

Db 230 TTCAAGGAGCTGGAGGCTGGGAGCCAGATGACGACCCCATAGAGAAACATAAAAGCAT 289

Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100

Db 290 TCGTCCGGTTCGCTTTCTTCTGTCAAGAGCAGTTTGAGAAATTAAACCTTGGTGA 349

Qy 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120

Db 350 TTTTGAAGAACTGGACAGAGAAAGCCCAAGAACAAAATTGCAAGAGAAACCAACAATAAG 409

Qy 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluLeuAlaAla 140

Db 410 AAGAAAGAAATTTGAGGAAACTGCGAAGAAAGTGGCGCGTCCCATCGACAGCTGGCTGCC 469

Qy 141 MetAsp 142

Db 470 ATGGAT 475

RESULT 4

US-10-172-118-566

Sequence 566, Application US/10172118

Publication No. US20030224374A1

GENERAL INFORMATION:

APPLICANT: Dai, Hongyue

APPLICANT: He, Yudong

APPLICANT: Linsley, Peter

APPLICANT: Mao, Mao

APPLICANT: Roberts, Chris

APPLICANT: Van 't Veer, Laura

APPLICANT: Bernards, Rene

TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

FILE REFERENCE: 9301-175-999

CURRENT APPLICATION NUMBER: US/10/172,118

CURRENT FILING DATE: 2002-06-14

PRIOR APPLICATION NUMBER: 60/380,770

PRIOR FILING DATE: 2002-05-14

NUMBER OF SEQ ID NOS: 2699

SEQ ID NO 566

LENGTH: 1619

TYPE: DNA

ORGANISM: Homo sapiens

PUBLICATION INFORMATION:

DATABASE ACCESSION NUMBER: NM\_001168

DATABASE ENTRY DATE: 2001-06-18

US-10-172-118-566

Alignment Scores:

Pred. No.: 2,54e-145 Length: 1619

Score: 142.00 Matches: 142

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-172-118-566 (1-1619)

Qy 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20

Db 50 ATGGGTGCCCGCAGCTTGCCTGCGCCCTTCTCAAGGACCAACCGCATCTCT 109

Qy 21 ThrPheLysAsnTrpPropheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40

Db 110 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGGCGCTGCACCCCGGAGCGGATGGCCGAG 169

Qy 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60

Db 170 GCTGGCTTCATCCACTGCCCTTCTTGGAGGGCTGGCGCTGCACCCCGGAGCGGATGGCCGAG 229

Qy 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80

Db 230 TTCAAGGAGCTGGAGGCTGGGAGCCAGATGACGACCCCATAGAGAAACATAAAAGCAT 289

Qy 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100

Db 290 TCGTCCGGTTGGCTTTCTCTCTCAAGAGCAGTTTGAAGAAATTAACCCCTTGGTGAA 349  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 Db 350 TTTTGAACCTGCAGAGAGAGCCCAAGACAAATTTGCAAGAGAAACCAACAATAAG 409  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
 Db 410 AAGAAGAAATTTGAGAACTCGGAAGAAAGTGCCCGTGCCATCGAGCAGCTGGCTGCC 469  
 QY 141 MetAsp 142  
 Db 470 ATGGAT 475

## RESULT 5

US-10-181-316-97  
 ; Sequence 97, Application US/10181316  
 ; Publication No. US20030211607A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: C. Frank Bennett  
 ; APPLICANT: Elizabeth J. Ackermann  
 ; APPLICANT: Eric B. Swayze  
 ; APPLICANT: Lex M. Cowsett  
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
 ; FILE REFERENCE: ISPH-0650  
 ; CURRENT APPLICATION NUMBER: US/10181,316  
 ; PRIOR FILING DATE: 2002-07-15  
 ; PRIOR APPLICATION NUMBER: PCT/US01/02939  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: 09/496,694  
 ; PRIOR FILING DATE: 2000-02-02  
 ; PRIOR APPLICATION NUMBER: 09/286,407  
 ; PRIOR FILING DATE: 1999-04-05  
 ; PRIOR APPLICATION NUMBER: 09/163,162  
 ; PRIOR FILING DATE: 1998-09-29  
 ; NUMBER OF SEQ ID NOS: 249  
 ; SEQ ID NO 97  
 ; LENGTH: 1619  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-181-316-97

Alignment Scores:  
 Pred. No.: 2,54e-145 Length: 1619  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-181-316-97 (1-1619)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 50 ATGGGTGCCCCGACGTTGCCCTTCTCTGGAGGCTGCGCCCTTCTCAAGGACCCGCACTCTCT 109  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 110 ACATTCAAGAACTGGCCCTTCTTGAGGGCTGCGCCCTGCAACCCCGGAGCGGATGGCCGAG 169  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 Db 170 GCTGGCTTCATCCACTGCCCACTGAGAACGAGCCAGACTTGGCCCACTGTTTCTCTGCG 229  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
 Db 230 TTCAAGGAGCTGGAAGCTGGAGCCAGATGACGACCCCATAGAGGAACATAAAGAGCAT 289  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluGluThrLeuGlyGlu 100  
 Db 290 TCGTCCGGTTGGCTTTCTCTCTCAAGAGCAGTTTGAAGAAATTAACCCCTTGGTGAA 349

QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 Db 350 TTTTGAACCTGCAGAGAGAGCCCAAGACAAATTTGCAAGAGAAACCAACAATAAG 409  
 QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaAla 140  
 Db 410 AAGAAGAAATTTGAGAACTCGGAAGAAAGTGCCCGTGCCATCGAGCAGCTGGCTGCC 469  
 QY 141 MetAsp 142  
 Db 470 ATGGAT 475

## RESULT 6

US-10-388-360-324  
 ; Sequence 324, Application US/10388360  
 ; Publication No. US20030225528A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GENOMIC HEALTH  
 ; APPLICANT: Baker, Joffe B.  
 ; APPLICANT: Cronin, Maureen T.  
 ; APPLICANT: Kiefer, Michael C.  
 ; APPLICANT: Shak, Steve  
 ; APPLICANT: Walker, Michael Graham  
 ; TITLE OF INVENTION: GENE EXPRESSION PROFILING IN BIOPSIED TUMOR TISSUES  
 ; FILE REFERENCE: 39740-0001US  
 ; CURRENT APPLICATION NUMBER: US/10/388,360  
 ; CURRENT FILING DATE: 2003-03-12  
 ; PRIOR FILING DATE: 2002-09-18  
 ; PRIOR APPLICATION NUMBER: US 60/412,049  
 ; PRIOR FILING DATE: 2002-03-13  
 ; NUMBER OF SEQ ID NOS: 384  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 324  
 ; LENGTH: 1619  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-388-360-324

Alignment Scores:  
 Pred. No.: 2,54e-145 Length: 1619  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-388-360-324 (1-1619)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
 Db 50 ATGGGTGCCCCGACGTTGCCCTTCTCTGGAGGCTGCGCCCTTCTCAAGGACCCGCACTCTCT 109  
 QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
 Db 110 ACATTCAAGAACTGGCCCTTCTTGAGGGCTGCGCCCTGCAACCCCGGAGCGGATGGCCGAG 169  
 QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
 Db 170 GCTGGCTTCATCCACTGCCCACTGAGAACGAGCCAGACTTGGCCCACTGTTTCTCTGCG 229  
 QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80  
 Db 230 TTCAAGGAGCTGGAAGCTGGAGCCAGATGACGACCCCATAGAGGAACATAAAGAGCAT 289  
 QY 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluGluThrLeuGlyGlu 100  
 Db 290 TCGTCCGGTTGGCTTTCTCTCTCAAGAGCAGTTTGAAGAAATTAACCCCTTGGTGAA 349  
 QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
 Db 350 TTTTGAACCTGCAGAGAGAGCCCAAGACAAATTTGCAAGAGAAACCAACAATAAG 409

QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaA 140  
|||||  
Db 410 AAGAAGAAATTTGAGAAATCGAGAAAGTGGCCCGTGGCATGAGCAGCTGGCTGCC 469  
  
QY 141 MetAsp 142  
|||||  
Db 470 ATGGAT 475

## RESULT 7

US-10-283-975A-445  
; Sequence 445, Application US/10283975A  
; Publication No. US20040110792A1  
; GENERAL INFORMATION:  
; APPLICANT: Ortho-Clinical Diagnostics, Inc.  
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia  
; FILE REFERENCE: CDS 293 PCT  
; CURRENT APPLICATION NUMBER: US/10/283,975A  
; CURRENT FILING DATE: 2002-10-30  
; PRIOR APPLICATION NUMBER: 60/340,938  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/338,997  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/340,081  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/341,012  
; PRIOR FILING DATE: 2001-10-30  
; NUMBER OF SEQ ID NOS: 900  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 445  
; LENGTH: 1619  
; TYPE: DNA  
; ORGANISM: HUMAN  
US-10-283-975A-445

Alignment Scores:  
Pred. No.: 2,54e-145 Length: 1619  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0

US-09-690-825-34 (1-142) x US-10-283-975A-445 (1-1619)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
|||||  
Db 50 ATGGGTGCCCCGACGTGGCCCCCTGCTGGCAGCCCTTTCTCAAGGACCACCGCATCTCT 109  
  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
|||||  
Db 110 ACATTCAGAACTGGCCCTTCTTGGAGGCTGGCCCTGCACCCCGAGCGGATGGCCGAG 169  
  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
|||||  
Db 170 GCTGGCTTTCATCCACTGCCCTCTTGGAGGCTGGCCCTGCACCCCGAGCGGATGGCCGAG 229  
  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80  
|||||  
Db 230 TTCAGAGCTGGAAGCTGGAGCCAGATGACGCCCTCCATGAGGAAATTAAGCAT 289  
  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100  
|||||  
Db 290 TCGTCCGTTGCGCTTTCCTTCTGTCAAGAACAGTTTGAAGAATTAACCCCTTGGTGAA 349  
  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
|||||  
Db 350 TTTTGAACCTGACAGAAAGAGCCAGAAACAAAATTGCAAGGAAACCAACATAAG 409  
  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaA 140  
|||||  
Db 410 AAGAAGAAATTTGAGAAATCGAGAAAGTGGCCCGTGGCATGAGCAGCTGGCTGCC 469  
  
QY 141 MetAsp 142  
|||||

Db 470 ATGGAT 475  
|||||

## RESULT 8

US-10-071-766-109  
; Sequence 109, Application US/10071766  
; Publication No. US20020192678A1  
; GENERAL INFORMATION:  
; APPLICANT: Huel-Mei Chen  
; TITLE OF INVENTION: GENES EXPRESSED IN SENESECE  
; FILE REFERENCE: PA-0043 US  
; CURRENT APPLICATION NUMBER: US/10/071,766  
; CURRENT FILING DATE: 2002-02-07  
; NUMBER OF SEQ ID NOS: 144  
; SOFTWARE: PERL Program  
; SEQ ID NO 109  
; LENGTH: 2404  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020192678A1 251651.4  
; NAME/KEY: unsure  
; LOCATION: 710, 717, 724, 982-1007  
; OTHER INFORMATION: a, t, c, g, or other  
US-10-071-766-109

Alignment Scores:  
Pred. No.: 3,66e-145 Length: 2404  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-690-825-34 (1-142) x US-10-071-766-109 (1-2404)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
|||||  
Db 56 ATGGGTGCCCCGACGTGGCCCCCTGCTGGCAGCCCTTTCTCAAGGACCACCGCATCTCT 115  
  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
|||||  
Db 116 ACATTCAGAACTGGCCCTTCTTGGAGGCTGGCCCTGCACCCCGAGCGGATGGCCGAG 175  
  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
|||||  
Db 176 GCTGGCTTTCATCCACTGCCCTCTTGGAGGCTGGCCCTGCACCCCGAGCGGATGGCCGAG 235  
  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluHisLysLysHis 80  
|||||  
Db 236 TTCAGAGCTGGAAGCTGGAGCCAGATGACGCCCTCCATGAGGAAACATAAAAGCAT 295  
  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluLeuThrLeuGlyGlu 100  
|||||  
Db 286 TCGTCCGTTGCGCTTTCCTTCTGTCAAGAACAGTTTGAAGAATTAACCCCTTGGTGAA 355  
  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
|||||  
Db 356 TTTTGAACCTGACAGAAAGAGCCAGAAACAAAATTGCAAGGAAACCAACATAAG 415  
  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGlnLeuAlaA 140  
|||||  
Db 416 AAGAAGAAATTTGAGAAATCGAGAAAGTGGCCCGTGGCATGAGCAGCTGGCTGCC 475  
  
QY 141 MetAsp 142  
|||||  
Db 476 ATGGAT 481

## RESULT 9

US-10-084-817-285  
; Sequence 285, Application US/10084817  
; Publication No. US20030119009A1

```

; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 285
; LENGTH: 2404
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 251651.4
; NAME/KEY: unsure
; LOCATION: 710, 717, 724
; OTHER INFORMATION: a, t, c, g, or other
US-10-084-817-285

Alignment Scores:
Pred. No.: 3,66e-145 Length: 2404
Score: 142.00 Matches: 142
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-690-825-34 (1-142) x US-10-084-817-285 (1-2404)
QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
DB 56 ATGGGTGCCCCGAGGTGGCCCCCTGCTGGGAGGCCCTTTCTCAAGGACCCAGCATCTCT 115
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
DB 116 ACATTCAAGAACTGCCCCCTTTCTGAGGGCTGCGCTCCACCCCGAGCGGATGGCCGAG 175
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60
DB 176 GCTGCTTCATCCACTGCCCCCTTGAGAACGAGCCAGACTGGCCCCAGTGTTCCTTCGTC 235
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisLysLysHis 80
DB 236 TTCAAGGAGCTTGAAGGCTGGAGCCAGATGACGACCCCATAGAGGAACATAAAAGCAT 295
QY 81 SerSerGlyCysAlaPheLeuSerValLysGlnPheGluGluLeuThrLeuGlyGlu 100
DB 296 TCGTCCGGTTCGGTTCCTTCCTTCTGTCGAAGAGCAGTTTGAAGAAATTAACCCCTGGTGA 355
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120
DB 356 TTTTGAACCTGGACAGAGAAGAGCCAGAACAAATTTCAAGGAAACCAACATAAG 415
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAlaAla 140
DB 416 AAGAAAGAAATTTGAGAAATCGGAAGAAAGTGGCGGTGCCATCGACGAGCTGGCTGCC 475
QY 141 MetAsp 142
DB 476 ATGGAT 481

RESULT 10
US-10-305-720-918
; Sequence 918, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.

; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 918
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20040010136A1 752848
; NAME/KEY: unsure
; LOCATION: (1)... (740)
; OTHER INFORMATION: a, t, c, g, or other
US-10-305-720-918

Alignment Scores:
Pred. No.: 4,74e-92 Length: 740
Score: 93.00 Matches: 130
Percent Similarity: 98.48% Conservatives: 0
Best Local Similarity: 98.48% Mismatches: 1
Query Match: 65.49% Indels: 2
DB: 16 Gaps: 0

US-09-690-825-34 (1-142) x US-10-305-720-918 (1-740)
QY 12 PropheLeuLysAspHisArgIleSerThrPheLysAsnTrpPropheLeuGluGlyCys 31
DB 51 CCCCTTCTCAAGGACCCAGCATCTCTACATTCAGAACTGCGCCCTTCTTGGAGGGCTGC 110
QY 32 AlaCysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrGlu-AsnGly 51
DB 111 GCCTGCACCCCGAGCGATGGCCGAGCGTGGCTTCATCCACTGCCCACTGA-NAACGA 169
QY 51 uProAspLeuAlaGlnCysPhePheCysPheLysGluLeuGluGlyTrpGluProAspAs 71
DB 170 GCACAGACTTGGCCCACTGTTCTTCTGCTTCAAGAGAGCTGGAAGGCTGGGAGCCAGATGA 229
QY 71 pAspProIleGluGluHisLysLysHisSerSerGlyCysAlaPheLeuSerValLysLy 91
DB 230 CGACCCCATAGAGNACATTAAGAGCATTCGTCGGTTGCGCTTTCCTTCTGTCAAGAA 289
QY 91 sGlnPheGluGluLeuThrLeuGlyGluPheLeuLysLeuAspArgGluArgAlaLysAs 111
DB 290 GCAGTTTCAAGAAATTAACCCCTTGGTGAATTTTGAACCTGGACAGAGAAAGCCAAAGAA 349
QY 111 nLysIleAlaLysGluThrAsnAsnLysLysLysGluPheGluGluThrAlaLysLysVa 131
DB 350 CAAAATTCGAAGGAAACCAACAATAAGAAAGAAATTTGAGGAACCTGCAAGAAAGT 409
QY 131 lArgArgAlaIleGluGlnLeuAlaAlaMetAsp 142
DB 410 GCGCGTGGCCATCGAGCAGCTGGCTGCCATGGAT 443

RESULT 11
US-09-918-186A-10
; Sequence 10, Application US/09918186A
; Patent No. US20020137708A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Elizabeth J. Ackermann
; APPLICANT: Eric E. Swayze
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: ISPH-0585
; CURRENT APPLICATION NUMBER: US/09/918,186A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/496,694

```



```
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 09/286,407
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 09/163,162
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 250
; SEQ ID NO 10
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (109)...(531)
US-09-918-186A-10

Alignment Scores:
Pred. No.: 4.4e-35 Length: 955
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 28.87% Indels: 0
DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x US-09-918-186A-10 (1-955)
QY 31 CysAlaCysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrGluAsn 50
Db 199 TGGCCTGCACCCAGAGCGAATGGCGGAGGCTGGCTTCATCCACTGCCCTACCGAGAAC 258
QY 51 GluProAspLeuAlaGlnCysPhePheCysPheLysGluLeuGluGlyTrpGluProAsp 70
Db 259 GAGCCTGATTGGCCCACTGTTTTCTGCTTTAAGGAATTGGAAGCTGGGAACCCGAT 318
QY 71 Asp 71
Db 319 GAC 321

RESULT 12
US-10-181-316-10
; Sequence 10, Application US/10181316
; Publication No. US20030211607A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Elizabeth J. Ackermann
; APPLICANT: Eric E. Swayze
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: ISPH-0650
; CURRENT APPLICATION NUMBER: US/10/181,316
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: PCT/US01/02939
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/496,694
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 09/286,407
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 09/163,162
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 249
; SEQ ID NO 10
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (109)...(531)
US-10-181-316-10

Alignment Scores:
Pred. No.: 4.4e-35 Length: 955
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

US-09-690-825-34 (1-142) x US-10-181-316-10 (1-955)
QY 31 CysAlaCysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrGluAsn 50
Db 199 TGGCCTGCACCCAGAGCGAATGGCGGAGGCTGGCTTCATCCACTGCCCTACCGAGAAC 258
QY 51 GluProAspLeuAlaGlnCysPhePheCysPheLysGluLeuGluGlyTrpGluProAsp 70
Db 259 GAGCCTGATTGGCCCACTGTTTTCTGCTTTAAGGAATTGGAAGCTGGGAACCCGAT 318
QY 71 Asp 71
Db 319 GAC 321

RESULT 13
US-10-179-730-33
; Sequence 33, Application US/10179730
; Publication No. US20030129611A1
; GENERAL INFORMATION:
; APPLICANT: BAQ, GANG
; APPLICANT: TSOURKAS, ANDREW
; APPLICANT: XU, YANGQING
; TITLE OF INVENTION: DUAL RESONANCE ENERGY TRANSFER NUCLEIC ACID PROBES
; FILE REFERENCE: 17625-0037
; CURRENT APPLICATION NUMBER: US/10/179,730
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/300,672
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: 60/303,258
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-179-730-33

Alignment Scores:
Pred. No.: 8.09e-35 Length: 121
Score: 40.00 Matches: 40
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 28.17% Indels: 0
DB: 15 Gaps: 0

US-09-690-825-34 (1-142) x US-10-179-730-33 (1-121)
QY 1 MetGlyAlaProThrLeuProProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20
Db 1 ATGGGTGCCCGACGTTGCCCTGCTGGCAGCCCTTCTCAAGGACCCACCGCATCTCT 60
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40
Db 61 ACATTCAGAACTGGCCCTTCTTGGAGGGTGGCCCTGCGACCCCGAGCGATGGCCGAG 120

RESULT 14
US-10-181-316-3
; Sequence 3, Application US/10181316
; Publication No. US20030211607A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Elizabeth J. Ackermann
; APPLICANT: Eric E. Swayze
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: ISPH-0650
; CURRENT APPLICATION NUMBER: US/10/181,316
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: PCT/US01/02939
```

, PRIOR FILING DATE: 2001-01-30  
 , PRIOR APPLICATION NUMBER: 09/496,694  
 , PRIOR FILING DATE: 2000-02-02  
 , PRIOR APPLICATION NUMBER: 09/286,407  
 , PRIOR FILING DATE: 1999-04-05  
 , PRIOR APPLICATION NUMBER: 09/163,162  
 , PRIOR FILING DATE: 1998-09-29  
 , NUMBER OF SEQ ID NOS: 349  
 , SEQ ID NO 3

```

; LENGTH: 14795
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2811)...(2921)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3174)...(3283)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5158)...(5275)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11955)...(12044)
US-10-181-316-3

Alignment Scores:
Pred. No.:      8 5e-32
Score:          39.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:    27.46%
DB:             13
Length:         14795
Matches:        39
Conservative:   0
Mismatches:    0
Indels:         0
Gaps:          0

```

US-09-690-825-34 (1-142) x US-10-181-316-3 (1-14795)

75	GlutHistisLysLysHisSerSerGlyCysAlaPheLeuSerValLysGlyGlnPheGlu	94
5158	GAGAAACATAAAGACATTCGTCGGTTCGCTTCTCTGTCGAAGACATTTGAA	5217
95	GluLeuThrLeuGlyGluPheLeuLysLeuAspArgGluArgAlaLysAsnLysIle	113
5218	CAATTACCCCTTGCGTAATTTTGAACCTGGACAGAGAAAGAGCCAGAAACAAAT	5274

```

; PRIOR APPLICATION NUMBER: US/60/2335, 863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 973
; LENGTH: 14796
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-954-456-973

```

Alignment Scores:		
Pred. No.:	8.5e-32	Length:
Score:	39.00	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	27.46%	Indels:
DB:	9	Gaps:

US-09-690-825-34 (1-142) x US-09-954-456-973 (1-14796)

QY	75	GluGluHisLysLysHisSerSerGlyCYsAlaPheLeuSerValLysLysGlnPheGlu	94
			21
Db	5159	GAGGACATAAAAGCATTCGCGCGTTCGCTTCTCTCAGAGCAGCAGTTTGA	5218
QY	95	GluLeuThrLeuGlyGlyPheLeuLysLeuAspArgGluArgAlaLysAsnLysIle	113
Db	5219	GAATTAACTCTTGTCGTAATTTTGAACTGTGACAGAGAAGAGCCAGAACAAAAT	5275

Alignment Scores:		
Pred. No.:	8.5e-32	Length: 14796
Score:	39.00	Matches: 39
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	27.46%	Indels: 0

```

DB:                                     9                               Caps:                                0

US-09-690-825-34 (1-142) x US-09-954-456-1636 (1-14796)

QY      75 GluGluHisLysLysHisSerSerGlyCysAlaPheLeuSerVallysLysGlnPheGlu 94
DB      5159 GAGGAACATATAAAGACATTCGTCGGTGGCTTTCTCTTCTGCAAGAAGCAGTTTGAA 5218

QY      95 GluLeuThrLeuGlyGluPheLeuLysLeuAspArgGluArgAlaLysAsnLysIle 113
DB      5219 GAATTAACCCCTGGTGAATTTTGGAACTGGACAGAGAAGAGCCAGCAAAATT 5275

RESULT 17
US-09-918-186A-3
; Sequence 3, Application US/09918186A
; Patent No. US20020137708A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Elizabeth J. Ackermann
; APPLICANT: Eric B. Sawvze
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: ISPH-0585
; CURRENT APPLICATION NUMBER: US/09/918,186A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/496,694
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 09/286,407
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 09/163,162
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 250
; SEQ ID NO 3
; LENGTH: 14796
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2811)...(2921)
; NAME/KEY: CDS
; LOCATION: (3174)...(3283)
; NAME/KEY: CDS
; LOCATION: (5158)...(5275)
; NAME/KEY: CDS
; LOCATION: (11955)...(12044)
US-09-918-186A-3

Alignment Scores:
Pred. No.:      8,58-32      Length:      14796
Score:          39.00      Matches:      39
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      27.46%      Indels:      0
DB:              9      Gaps:      0

US-09-690-825-34 (1-142) x US-09-918-186A-3 (1-14796)

QY      75 GluGluHisLysLysHisSerSerGlyCysAlaPheLeuSerVallysLysGlnPheGlu 94
DB      5159 GAGGAACATATAAAGACATTCGTCGGTGGCTTTCTCTTCTGCAAGAAGCAGTTTGAA 5218

QY      95 GluLeuThrLeuGlyGluPheLeuLysLeuAspArgGluArgAlaLysAsnLysIle 113
DB      5219 GAATTAACCCCTGGTGAATTTTGGAACTGGACAGAGAAGAGCCAGCAAAATT 5275

RESULT 18
US-09-880-107-3421
; Sequence 3421, Application US/09980107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe

```

Query Match: 27.46% Indels: 0  
DB: 12 Gaps: 0  
US-09-690-825-34 (1-142) x US-09-968-007A-522 (1-14796)

Qy 75 GluGluHisLysLysHisSerGlyCysAlaPheLeuSerValLysLysGlnPheGlu 94  
Db 5159 GAGGAACATAAAAGCATTCGCGGTTGCCCTTCTCTGCAAGAGCAGTTTGAA 5218  
Qy 95 GluLeuThrLeuGlyGluPheLeuLysLeuAspArgGluArgAlaLysAsnLysile 113  
Db 5219 GAATTAACCCCTGGTGAATTTTGAACACTGGACAGAGAAAGAGCAACAAATT 5275

RESULT 20  
US-10-138-618-35  
; Sequence 35, Application US/10138618  
; Publication No. US20030100525A1  
; GENERAL INFORMATION:  
; APPLICANT: Altieri, Dario C.  
; TITLE OF INVENTION: CELLULAR APOPTOSIS, AND ITS MODULATION  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
; STREET: 1800 M Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036-5869

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10138,618  
; FILING DATE: 06-May-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/975,080  
; FILING DATE: 20-NOV-1997  
; APPLICATION NUMBER: US 60/031,435  
; FILING DATE: 20-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adler, Reid G.  
; REGISTRATION NUMBER: 30,988  
; REFERENCE/DOCKET NUMBER: 044574-5022-01-WO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-467-7000  
; TELEFAX: 202-467-7176

INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14796 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
US-10-138-618-35

Alignment Scores:  
Pred. No.: 8.5e-32 Length: 14796  
Score: 39.00 Matches: 39  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 27.46% Indels: 0  
DB: 15 Gaps: 0  
US-09-690-825-34 (1-142) x US-10-138-618-35 (1-14796)

Qy 75 GluGluHisLysLysHisSerGlyCysAlaPheLeuSerValLysLysGlnPheGlu 94  
Db 5159 GAGGAACATAAAAGCATTCGCGGTTGCCCTTCTCTGCAAGAGCAGTTTGAA 5218

Qy 95 GluLeuThrLeuGlyGluPheLeuLysLeuAspArgGluArgAlaLysAsnLysile 113  
Db 5219 GAATTAACCCCTGGTGAATTTTGAACACTGGACAGAGAAAGAGCAACAAATT 5275

RESULT 21  
US-10-029-386-23672/c  
; Sequence 23672, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 23672  
; LENGTH: 120  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR17.3  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.66  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.52  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2  
; OTHER INFORMATION: NT HIT: U75285.1, EVALUE 2.00e-61  
; OTHER INFORMATION: EST HUMAN HIT: AA994786.1, EVALUE 5.00e-60  
; OTHER INFORMATION: SWISSPROT HIT: O15392, EVALUE 2.00e-18  
US-10-029-386-23672

Alignment Scores:  
Pred. No.: 1.9e-30 Length: 120  
Score: 36.00 Matches: 36  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 25.35% Indels: 0  
DB: 15 Gaps: 0  
US-09-690-825-34 (1-142) x US-10-029-386-23672 (1-120)

Qy 38 MetaGluAlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCys 57  
Db 117 ATGGCCGAGGCTGGCTTCATCCACTGCCCACTGAGACGAGCCAGACTTGGCCCATGT 58  
Qy 58 PhePheCysPheLysGluLeuGlyTrpGluProAspAspPro 73  
Db 57 TTCTTCTGCTTCAAGAGCTGGAAGGCTGGAGCCAGATGACGACCCC 10

RESULT 22  
US-10-029-386-9972/c  
; Sequence 9972, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 9972  
; LENGTH: 555  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR17.3

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.66  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2  
OTHER INFORMATION: EST\_HUMAN HIT: AA954786.1, EVALUATE 1.00e-122  
OTHER INFORMATION: SWISSPROT HIT: O15392, EVALUATE 4.00e-19  
OTHER INFORMATION: NT HIT: U75285.1, EVALUATE 0.00e+00  
NAME/KEY: unsure  
LOCATION: 320  
US-10-029-386-9972

Alignment Scores:  
Pred. No.: 7,82e-30 Length: 555  
Score: 36.00 Matches: 36  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 25.35% Indels: 0  
DB: 15 Gaps: 0

US-09-690-825-34 (1-142) x US-10-029-386-9972 (1-555)

QY 38 MetalAlaGluAlaGlyPheLeuHisCysProThrGluAsnGluProAspLeuAlaGlnCys 57  
DB 222 ATGGCGAGGCTGGCTTCATCCACTGCCCACTGAGAACGAGCCAGACTTGGCCAGTGT 163  
QY 58 PhePheCysPheGluLeuGluGlyTrpGluProAspAspPro 73  
DB 162 TTCCTTCTGCTCAAGAGAGCTGGAAGCTGGGAGCCAGATGAGACCCC 115

RESULT 23

US-10-108-877-2

Sequence 2, Application US/10108877  
Publication No. US20030083482A1  
GENERAL INFORMATION:  
APPLICANT: Fox Chase Cancer Center  
TITLE OF INVENTION: Compositions and Methods for  
TITLE OF INVENTION: P53-Mediated Repression of Gene Expression  
FILE REFERENCE: FCCC 99-03  
CURRENT APPLICATION NUMBER: US/10/108,877  
PRIOR FILING DATE: 2002-03-28  
PRIOR FILING DATE: PCT/US00/27078  
PRIOR FILING DATE: 2000-10-02  
PRIOR FILING DATE: 60/157,171  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 1100  
TYPE: DNA  
ORGANISM: murine  
US-10-108-877-2

Alignment Scores:  
Pred. No.: 1,47e-29 Length: 1100  
Score: 36.00 Matches: 36  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 25.35% Indels: 0  
DB: 15 Gaps: 0

US-09-690-825-34 (1-142) x US-10-108-877-2 (1-1100)

QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
DB 991 ATGGGTGCCCCGAGCTGTGCCCTGCTGGCAGCCCTTCTCAAGACCCAGCATCTCT 1050  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGlu 36  
DB 1051 ACATTCAAGAACTGGCCCTTCTTGGAGGCTGGCTGCACCCCGGAG 1098

RESULT 24

US-09-833-381-1331

Sequence 1331, Application US/09833381  
Patent No. US20020132090A1  
GENERAL INFORMATION:  
APPLICANT: Robison, Keith E.  
TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs  
FILE REFERENCE: 5800-119  
CURRENT APPLICATION NUMBER: US/09/833,381  
CURRENT FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: 09/516,448  
PRIOR FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 2050  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1331  
LENGTH: 332  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)-(332)  
OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1331

Alignment Scores:  
Pred. No.: 5,57e-06 Length: 332  
Score: 14.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.86% Indels: 0  
DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x US-09-833-381-1331 (1-332)

QY 68 GluProAspAspProThrGluGluHisLysLysHisSer 81  
DB 240 GAGCCAGATGACGCCCNATAGAGACATATAAAGCATTCG 281

RESULT 25

US-10-400-382-265  
Sequence 265, Application US/10400382  
Publication No. US20030190659A1  
GENERAL INFORMATION:  
APPLICANT: Lacasse, Eric  
APPLICANT: McManus, Daniel  
APPLICANT: Durkin, Jonathan P.  
TITLE OF INVENTION: Antisense IAP Nucleobase Oligomers and  
TITLE OF INVENTION: Uses Thereof  
FILE REFERENCE: 07891/025004  
CURRENT APPLICATION NUMBER: US/10/400,382  
CURRENT FILING DATE: 2003-03-27  
PRIOR APPLICATION NUMBER: US 60/367,853  
PRIOR FILING DATE: 2002-03-27  
NUMBER OF SEQ ID NOS: 460  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 265  
LENGTH: 30  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: based on Homo sapiens.  
US-10-400-382-265

Alignment Scores:  
Pred. No.: 0.177 Length: 30  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.34% Indels: 0  
DB: 15 Gaps: 0

US-09-690-825-34 (1-142) x US-10-400-382-265 (1-30)

QY 69 ProAspAspProThrGluGluHis 77  
|||||

Db 3 CCAGATGACGACCCCATAGAGGAACAT 29

## RESULT 26

US-10-636-065-199  
; Sequence 199, Application US/10636065  
; Publication No. US20040127694A1  
; GENERAL INFORMATION:  
; APPLICANT: Korneluk, Robert G.  
; APPLICANT: LaCasse, Eric  
; APPLICANT: Baird, Stephen  
; APPLICANT: Holcik, Martin  
; APPLICANT: Young, Sean  
; TITLE OF INVENTION: Antisense IAP Nucleic Acids and Uses  
; FILE REFERENCE: 07891/025005  
; CURRENT APPLICATION NUMBER: US/10/636,065  
; CURRENT FILING DATE: 2003-08-07  
; PRIOR APPLICATION NUMBER: 09/672,717  
; PRIOR FILING DATE: 2000-09-28  
; NUMBER OF SEQ ID NOS: 231  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 199  
; LENGTH: 30  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: based on Homo sapiens  
US-10-636-065-199

Alignment Scores:  
Pred. No.: 0.177 Length: 30  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.34% Indels: 0  
DB: 17 Gaps: 0

US-09-690-825-34 (1-142) x US-10-636-065-199 (1-30)

Qy 69 ProAspAspProfileGluGluHis 77

Db 3 CCAGATGACGACCCCATAGAGGAACAT 29

## RESULT 27

US-09-918-186A-13  
; Sequence 13, Application US/09918186A  
; Patent No. US20020137708A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Elizabeth J. Ackermann  
; APPLICANT: Eric E. Swayze  
; APPLICANT: Lex M. Cowsert  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
; FILE REFERENCE: ISFH-0585  
; CURRENT APPLICATION NUMBER: US/09/918,186A  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 09/496,694  
; PRIOR FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 09/286,407  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 09/163,162  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 250  
; SEQ ID NO 13  
; LENGTH: 26  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: PCR Probe  
US-09-918-186A-13

Alignment Scores:  
Pred. No.: 1.93 Length: 26

Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x US-09-918-186A-13 (1-26)

Qy 62 LysGluLeuGluGlyTyrGluPro 69

Db 2 AAGGAATTGGAAGGCTGGGAACCC 25

## RESULT 28

US-10-181-316-13  
; Sequence 13, Application US/10181316  
; Publication No. US20030211607A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Elizabeth J. Ackermann  
; APPLICANT: Eric E. Swayze  
; APPLICANT: Lex M. Cowsert  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
; FILE REFERENCE: ISFH-0650  
; CURRENT APPLICATION NUMBER: US/10/181,316  
; CURRENT FILING DATE: 2002-07-16  
; PRIOR APPLICATION NUMBER: PCT/US01/02939  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 09/496,694  
; PRIOR FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 09/286,407  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 09/163,162  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 249  
; SEQ ID NO 13  
; LENGTH: 26  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: PCR Probe  
US-10-181-316-13

Alignment Scores:  
Pred. No.: 1.93 Length: 26  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-181-316-13 (1-26)

Qy 62 LysGluLeuGluGlyTyrGluPro 69

Db 2 AAGGAATTGGAAGGCTGGGAACCC 25

## RESULT 29

US-09-833-381-1330  
; Sequence 1330, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1330  
; LENGTH: 281  
; TYPE: DNA

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(281)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1330

Alignment Scores:
Pred. No.: 17.4 Length: 281
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 9 Gaps: 0

US-09-690-825-34 (1-142) x US-09-833-381-1330 (1-281)

QY 60 CysPheLysGluLeuGluGlyTrrp 67
|||||
Db 258 TGCCTCAAGGAGCTGGAGGCTGG 281

RESULT 30
US-10-027-632-65868
; Sequence 65868, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65868
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-65868

Alignment Scores:
Pred. No.: 25.2 Length: 420
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-027-632-65868 (1-420)

QY 79 LysHisSerGlyCysAlaphe 86
|||||
Db 21 AAACATAGTTCAGGCTGTCATT 44

RESULT 31
US-10-027-632-65868
; Sequence 65868, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65868
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-65868

Alignment Scores:
Pred. No.: 25.2 Length: 420
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-027-632-65868 (1-420)

QY 79 LysHisSerGlyCysAlaphe 86
|||||
Db 21 AAACATAGTTCAGGCTGTCATT 44

RESULT 31
US-10-027-632-65868
; Sequence 65868, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65869
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-65869

Alignment Scores:
Pred. No.: 25.2 Length: 420
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-027-632-65869 (1-420)

QY 79 LysHisSerGlyCysAlaphe 86
|||||
Db 21 AAACATAGTTCAGGCTGTCATT 44

RESULT 32
US-10-027-632-65870
; Sequence 65870, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65870
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Human

```

```
US-10-027-632-65870
Alignment Scores:
Pred. No.: 25.2 Length: 420
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-027-632-65870 (1-420)
Qy 79 LysHisSerSerGlyCysAlaphe 86
Db 21 AAACATAGTTCAGGCTGTGCATT 44

RESULT 33
US-10-027-632-65868
; Sequence 65868, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65868
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-65868
Alignment Scores:
Pred. No.: 25.2 Length: 420
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 16 Gaps: 0

US-09-690-825-34 (1-142) x US-10-027-632-65868 (1-420)
Qy 79 LysHisSerSerGlyCysAlaphe 86
Db 21 AAACATAGTTCAGGCTGTGCATT 44

RESULT 34
US-10-027-632-65869
; Sequence 65869, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65869
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-65869
Alignment Scores:
Pred. No.: 25.2 Length: 420
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 16 Gaps: 0

US-09-690-825-34 (1-142) x US-10-027-632-65869 (1-420)
Qy 79 LysHisSerSerGlyCysAlaphe 86
Db 21 AAACATAGTTCAGGCTGTGCATT 44

RESULT 35
US-10-027-632-65870
; Sequence 65870, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65870
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-65870
Alignment Scores:
Pred. No.: 25.2 Length: 420
Score: 8.00 Matches: 8
```



Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 16 Gaps: 0  
US-09-690-825-34 (1-142) x US-10-027-632-65870 (1-420)

QY 79 LysHisSerGlyCysAlaPhe 86  
DB 21 AAACATAGTTCAGGCTGTGCATT 44

RESULT 36  
US-10-424-599-114486  
; Sequence 114486, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David X  
; APPLICANT: Zhou Yihua  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 114486  
; LENGTH: 454  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_7438C.1  
US-10-424-599-114486

Alignment Scores: 27.1 Length: 454  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 13 Gaps: 0  
US-09-690-825-34 (1-142) x US-10-424-599-114486 (1-454)

QY 2 GlyAlaProThrLeuProAla 9  
DB 271 GGTGCCCGGACCTTCCCGGCT 294

RESULT 37  
US-09-918-186A-231  
; Sequence 231, Application US/09918186A  
; Patent No. US20020137708A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Elizabeth J. Ackermann  
; APPLICANT: Eric B. Swayze  
; APPLICANT: Lex M. Cowsett  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
; FILE REFERENCE: ISPH-0385  
; CURRENT APPLICATION NUMBER: US/09/918,186A  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 09/496,694  
; PRIOR FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 09/286,407  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 09/163,162  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 250  
; SEQ ID NO 231  
; LENGTH: 515  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:

; NAME/KEY: unsure  
; LOCATION: 266  
; OTHER INFORMATION: unknown  
; NAME/KEY: CDS  
; LOCATION: (301)...(384)  
US-09-918-186A-231  
Alignment Scores: 30.4 Length: 515  
Pred. No.: 8.00 Matches: 8  
Score: 8.00 Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 5.63% Gaps: 0  
DB: 9  
US-09-690-825-34 (1-142) x US-09-918-186A-231 (1-515)  
QY 122 LysGluPheGluThrAlaLys 129  
DB 325 AAAGAGTTTGAGAGACTGCAAG 348  
RESULT 38  
US-10-181-316-231  
; Sequence 231, Application US/10181316  
; Publication No. US20030211607A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Elizabeth J. Ackermann  
; APPLICANT: Eric B. Swayze  
; APPLICANT: Lex M. Cowsett  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
; FILE REFERENCE: ISPH-0650  
; CURRENT APPLICATION NUMBER: US/10/181,316  
; CURRENT FILING DATE: 2002-07-16  
; PRIOR APPLICATION NUMBER: PCT/US01/02939  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 09/496,694  
; PRIOR FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 09/286,407  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 09/163,162  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 249  
; SEQ ID NO 231  
; LENGTH: 515  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 266  
; OTHER INFORMATION: unknown  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (301)...(384)  
US-10-181-316-231  
Alignment Scores: 30.4 Length: 515  
Pred. No.: 8.00 Matches: 8  
Score: 8.00 Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 5.63% Gaps: 0  
DB: 13  
US-09-690-825-34 (1-142) x US-10-181-316-231 (1-515)  
QY 122 LysGluPheGluThrAlaLys 129  
DB 325 AAAGAGTTTGAGAGACTGCAAG 348  
RESULT 39  
US-09-918-995-15078  
; Sequence 15078, Application US/09918995

Publication No. US20030073623A1  
 GENERAL INFORMATION:  
 APPLICANT: Hyseq, Inc.  
 TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
 FILE REFERENCE: 20411-756  
 CURRENT APPLICATION NUMBER: US/09/918,995  
 PRIOR FILING DATE: 2001-07-30  
 PRIOR APPLICATION NUMBER: US/09/235,076  
 PRIOR FILING DATE: 1999-01-20  
 NUMBER OF SEQ ID NOS: 38054  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 15078  
 LENGTH: 587  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (1)...(587)  
 OTHER INFORMATION: n = A,T,C or G  
 US-09-918-995-15078

Alignment Scores:  
 Pred. No.: 34.3 Length: 587  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.63% Indels: 0  
 DB: 10 Gaps: 0

US-09-690-825-34 (1-142) x US-09-918-995-15078 (1-587)

QY 1 MetGlyAlaProThrLeuProPro 8  
 DB 367 ATGGAGGCCCCACACTCCCTCCT 390

## RESULT 40

US-10-369-493-35886/c  
 Sequence 35886, Application US/10369493  
 Publication No. US20030233675A1  
 GENERAL INFORMATION:  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Hinkle, Gregory J.  
 APPLICANT: Slater, Steven C.  
 APPLICANT: Goldman, Barry S.  
 APPLICANT: Chen, Xianfeng  
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 FILE REFERENCE: 38-10(52052)B  
 CURRENT APPLICATION NUMBER: US/10/369,493  
 PRIOR FILING DATE: 2003-02-28  
 PRIOR APPLICATION NUMBER: US 60/360,039  
 PRIOR FILING DATE: 2002-02-21  
 NUMBER OF SEQ ID NOS: 47374  
 SEQ ID NO 35886  
 LENGTH: 1224  
 TYPE: DNA  
 ORGANISM: Mesorhizobium loti  
 US-10-369-493-35886

Alignment Scores:  
 Pred. No.: 67.6 Length: 1224  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.63% Indels: 0  
 DB: 16 Gaps: 0

US-09-690-825-34 (1-142) x US-10-369-493-35886 (1-1224)

QY 94 GluGluLeuThrLeuGlyGluPhe 101  
 DB 351 GAGGAGTTCAGCTGGGTGAATTT 328

RESULT 41  
 US-10-050-704-14  
 Sequence 14, Application US/10050704  
 Publication No. US20030050442A1  
 GENERAL INFORMATION:  
 APPLICANT: Ruben et al.  
 TITLE OF INVENTION: 62 Human Secreted Proteins  
 FILE REFERENCE: P2039P1  
 CURRENT APPLICATION NUMBER: US/10/050,704  
 CURRENT FILING DATE: 2002-01-18  
 PRIOR APPLICATION NUMBER: 09/684,524  
 PRIOR FILING DATE: 2000-10-10  
 PRIOR APPLICATION NUMBER: PCT/US00/08979  
 PRIOR FILING DATE: 2000-04-06  
 PRIOR APPLICATION NUMBER: 60/128,693  
 PRIOR FILING DATE: 1999-04-09  
 PRIOR APPLICATION NUMBER: 60/130,991  
 PRIOR FILING DATE: 1999-04-26  
 NUMBER OF SEQ ID NOS: 344  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 14  
 LENGTH: 1441  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: SITE  
 LOCATION: (1436)  
 OTHER INFORMATION: n equals a,t,g, or c  
 NAME/KEY: SITE  
 LOCATION: (1438)  
 OTHER INFORMATION: n equals a,t,g, or c  
 US-10-050-704-14

Alignment Scores:  
 Pred. No.: 78.7 Length: 1441  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.63% Indels: 0  
 DB: 15 Gaps: 0

US-09-690-825-34 (1-142) x US-10-050-704-14 (1-1441)

QY 1 MetGlyAlaProThrLeuProPro 8  
 DB 1343 ATGGAGGCCCCACACTCCCTCCT 1366

## RESULT 42

US-10-798-512-14  
 Sequence 14, Application US/10798512  
 Publication No. US20040152164A1  
 GENERAL INFORMATION:  
 APPLICANT: Ruben et al.  
 TITLE OF INVENTION: 62 Human Secreted Proteins  
 FILE REFERENCE: P2039P1  
 CURRENT APPLICATION NUMBER: US/10/798,512  
 CURRENT FILING DATE: 2004-03-12  
 PRIOR APPLICATION NUMBER: US/09/684,524  
 PRIOR FILING DATE: 2000-10-10  
 PRIOR APPLICATION NUMBER: PCT/US00/08979  
 PRIOR FILING DATE: 2000-04-06  
 PRIOR APPLICATION NUMBER: 60/128,693  
 PRIOR FILING DATE: 1999-04-09  
 PRIOR APPLICATION NUMBER: 60/130,991  
 PRIOR FILING DATE: 1999-04-26  
 NUMBER OF SEQ ID NOS: 344  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 14  
 LENGTH: 1441  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:

```

; NAME/KEY: SITE
; LOCATION: (1436)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1438)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-798-512-14

Alignment Scores:
Pred. No.: 78.7 Length: 1441
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 17 Gaps: 0

US-09-690-825-34 (1-142) x US-10-798-512-14 (1-1441)
QY 1 MetGlyAlaProThreuProPro 8
Db 1343 ATGGGAGCCCCACACTCCCTCT 1366

RESULT 43
US-09-946-374-51
; Sequence 51, Application US/09946374
; Publication No. US20030073129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William L.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: F2830P1C1
; CURRENT APPLICATION NUMBER: US/09/946,374
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099602
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099642
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099741
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099754
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099763
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099792
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099808
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099815
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100388
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100584
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100661
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100662
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100664
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100683
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100684
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100710
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100711
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100849
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100930
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101014
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101068
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101071
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101279
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: 60/101471
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101472
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101474
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101475
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101476
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23

```

; PRIOR APPLICATION NUMBER: 60/101479  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101738  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101741  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101743  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101915  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101916  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/102207  
; PRIOR FILING DATE: 1998-09-29  
; PRIOR APPLICATION NUMBER: 60/102240  
; PRIOR FILING DATE: 1998-09-29  
; PRIOR APPLICATION NUMBER: 60/102307  
; PRIOR FILING DATE: 1998-09-29  
; PRIOR APPLICATION NUMBER: 60/102330  
; PRIOR FILING DATE: 1998-09-29  
; PRIOR APPLICATION NUMBER: 60/102331  
; PRIOR FILING DATE: 1998-09-29  
; PRIOR APPLICATION NUMBER: 60/102484  
; PRIOR FILING DATE: 1998-09-30  
; PRIOR APPLICATION NUMBER: 60/102487  
; PRIOR FILING DATE: 1998-09-30  
; PRIOR APPLICATION NUMBER: 60/102570  
; PRIOR FILING DATE: 1998-09-30  
; PRIOR APPLICATION NUMBER: 60/102571  
; PRIOR FILING DATE: 1998-09-30  
; PRIOR APPLICATION NUMBER: 60/102684  
; PRIOR FILING DATE: 1998-10-01  
; PRIOR APPLICATION NUMBER: 60/102687  
; PRIOR FILING DATE: 1998-10-01  
; PRIOR APPLICATION NUMBER: 60/102965  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 60/103258  
; PRIOR FILING DATE: 1998-10-06  
; PRIOR APPLICATION NUMBER: 60/103314  
; PRIOR FILING DATE: 1998-10-07  
; PRIOR APPLICATION NUMBER: 60/103315  
; PRIOR FILING DATE: 1998-10-07  
; PRIOR APPLICATION NUMBER: 60/103328  
; PRIOR FILING DATE: 1998-10-07  
; PRIOR APPLICATION NUMBER: 60/103395  
; PRIOR FILING DATE: 1998-10-07  
; PRIOR APPLICATION NUMBER: 60/103396  
; PRIOR FILING DATE: 1998-10-07  
; PRIOR APPLICATION NUMBER: 60/103401  
; PRIOR FILING DATE: 1998-10-07  
; PRIOR APPLICATION NUMBER: 60/103449  
; PRIOR FILING DATE: 1998-10-06  
; PRIOR APPLICATION NUMBER: 60/103633  
; PRIOR FILING DATE: 1998-10-08  
; PRIOR APPLICATION NUMBER: 60/103678  
; PRIOR FILING DATE: 1998-10-08  
; PRIOR APPLICATION NUMBER: 60/103679  
; PRIOR FILING DATE: 1998-10-08  
; PRIOR APPLICATION NUMBER: 60/103711  
; PRIOR FILING DATE: 1998-10-08  
; PRIOR APPLICATION NUMBER: 60/104257  
; PRIOR FILING DATE: 1998-10-14  
; PRIOR APPLICATION NUMBER: 60/104987  
; PRIOR FILING DATE: 1998-10-20  
; PRIOR APPLICATION NUMBER: 60/105000  
; PRIOR FILING DATE: 1998-10-20  
; PRIOR APPLICATION NUMBER: 60/105002  
; PRIOR FILING DATE: 1998-10-20  
; PRIOR APPLICATION NUMBER: 60/105104  
; PRIOR FILING DATE: 1998-10-21  
; PRIOR APPLICATION NUMBER: 60/105169  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: 60/105266

; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: 60/105693  
; PRIOR FILING DATE: 1998-10-26  
; PRIOR APPLICATION NUMBER: 60/105694  
; PRIOR FILING DATE: 1998-10-26  
; PRIOR APPLICATION NUMBER: 60/105807  
Alignment Scores:  
Pred. No.: 93.3 Length: 1734  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 10 Gaps: 0  
US-09-690-825-34 (1-142) x US-09-946-374-51 (1-1734)  
QY 1 MetGlyAlaProThrLeuProPro 8  
DB 1534 ATGGAGCCCCCACACTCCCTCCT 1557  
RESULT 44  
US-10-015-395A-51  
; Sequence 51, Application US/10015395A  
; Publication No. US20040073015A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas P.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2830PlC57  
; CURRENT APPLICATION NUMBER: US/10/015,395A  
; CURRENT FILING DATE: 2001-12-12  
; Prior application removed - See file Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 477  
; SEQ ID NO 51  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-015-395A-51  
Alignment Scores:  
Pred. No.: 93.3 Length: 1734  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 12 Gaps: 0  
US-09-690-825-34 (1-142) x US-10-015-395A-51 (1-1734)  
QY 1 MetGlyAlaProThrLeuProPro 8  
DB 1534 ATGGAGCCCCCACACTCCCTCCT 1557  
RESULT 45  
US-10-206-915-201  
; Sequence 201, Application US/10206915  
; Publication No. US20040029221A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C464

CURRENT APPLICATION NUMBER: US/10/206,915

PRIOR FILING DATE: 2002-07-26

PRIOR APPLICATION NUMBER: 10/052586

PRIOR FILING DATE: 2002-01-15

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059266

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/063120

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063121

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063486

PRIOR FILING DATE: 1997-10-21

PRIOR APPLICATION NUMBER: 60/063540

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063541

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

US-10-206-915-201

Alignment Scores:

Pred. No.: 93.3 Length: 1734  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-206-915-201 (1-1734)

Qy 1 MetGlyAlaProThrLeuProPro 8

Db 1534 ATGGAGCCCCCACACTCCCTCT 1557

RESULT 46

US-10-199-670-201

Sequence 201, Application US/10199670

Publication No. US20040033560A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C464

CURRENT APPLICATION NUMBER: US/10/201,858

PRIOR FILING DATE: 2002-07-23

PRIOR APPLICATION NUMBER: 10/052586

PRIOR FILING DATE: 2002-01-15

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059266

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C401

CURRENT APPLICATION NUMBER: US/10/199,670

PRIOR FILING DATE: 2002-07-19

PRIOR APPLICATION NUMBER: 10/052586

PRIOR FILING DATE: 2002-01-15

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059266

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/063120

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063121

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063486

PRIOR FILING DATE: 1997-10-21

PRIOR APPLICATION NUMBER: 60/063540

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063541

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 201

LENGTH: 1734

TYPE: DNA

ORGANISM: Homo Sapien

US-10-199-670-201

Alignment Scores:

Pred. No.: 93.3 Length: 1734  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-199-670-201 (1-1734)

Qy 1 MetGlyAlaProThrLeuProPro 8

Db 1534 ATGGAGCCCCCACACTCCCTCT 1557

RESULT 47

US-10-201-858-201

Sequence 201, Application US/10201858

Publication No. US2004003837A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C464

CURRENT APPLICATION NUMBER: US/10/201,858

PRIOR FILING DATE: 2002-07-23

PRIOR APPLICATION NUMBER: 10/052586

PRIOR FILING DATE: 2002-01-15

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059266

```

; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 201
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-201-858-201

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-201-858-201 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8
DB 1534 ATGGAGGCCCCACACTCCCTCT 1557

RESULT 48
US-10-081-056-149
; Sequence 149, Application US/10081056
; Publication No. US20040043927A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerriksen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Williams, P.Mickey
; APPLICANT: Ye, Weilan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS
; FILE REFERENCE: F3235P1C1
; CURRENT APPLICATION NUMBER: US/10/081,056
; CURRENT FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: PCT/US01/21735
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/219,556
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/220,624
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/220,664
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: 2000-07-28

```

```

; PRIOR APPLICATION NUMBER: US 60/222,695
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US 09/643,657
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/23522
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/230,978
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/000,000
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 09/664,610
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/242,922
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 09/709,238
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/30952
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/30873
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/747,259
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/34956
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: US 09/767,609
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: US 09/796,498
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06666
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: US 09/802,706
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 09/808,689
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 09/816,744
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 09/828,366
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/854,208
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 09/854,280
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 09/866,028
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 09/866,034
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: PCT/US01/17092
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 09/870,574
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US01/17443
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US01/17800
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US01/19692
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: PCT/US01/00000
; PRIOR FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 383
; SEQ ID NO 149
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homosapiens
US-10-081-056-149

```

Alignment Scores:  
 Pred. No.: 93.3 Length: 1734  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.63% Indels: 0  
 DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-081-056-149 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8  
 DB 1534 ATGGAGCCCCCACACTCCCTCT 1557

RESULT 49

US-10-219-535-155  
 ; Sequence 155, Application US/10219555  
 ; Publication No. US2004004179A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Gerritsen, Mary  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stephan, Jean-Philippe F.  
 ; APPLICANT: Watanabe, Colin L.  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3530P1C103  
 ; CURRENT APPLICATION NUMBER: US/10/219,535  
 ; CURRENT FILING DATE: 2002-08-14  
 ; PRIOR APPLICATION NUMBER: 10/119,480  
 ; PRIOR FILING DATE: 2002-04-09  
 ; PRIOR APPLICATION NUMBER: 60/059113  
 ; PRIOR FILING DATE: 1997-09-17  
 ; PRIOR APPLICATION NUMBER: 60/062287  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/063549  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/064103  
 ; PRIOR FILING DATE: 1997-10-31  
 ; PRIOR APPLICATION NUMBER: 60/069873  
 ; PRIOR FILING DATE: 1997-12-17  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/079294  
 ; PRIOR FILING DATE: 1998-03-25  
 ; PRIOR APPLICATION NUMBER: 60/079656  
 ; PRIOR FILING DATE: 1998-03-26  
 ; PRIOR APPLICATION NUMBER: 60/079728  
 ; PRIOR FILING DATE: 1998-03-27  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 246  
 ; SEQ ID NO 155  
 ; LENGTH: 1734  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-219-535-155

Alignment Scores:  
 Pred. No.: 93.3 Length: 1734  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.63% Indels: 0  
 DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-219-535-155 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8  
 DB 1534 ATGGAGCCCCCACACTCCCTCT 1557

RESULT 50

US-10-232-230-155  
 ; Sequence 155, Application US/10232230  
 ; Publication No. US2004004180A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Gerritsen, Mary  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stephan, Jean-Philippe F.  
 ; APPLICANT: Watanabe, Colin L.  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3530P1C103  
 ; CURRENT APPLICATION NUMBER: US/10/232,230  
 ; CURRENT FILING DATE: 2002-08-29  
 ; PRIOR APPLICATION NUMBER: 10/119,480  
 ; PRIOR FILING DATE: 2002-04-09  
 ; PRIOR APPLICATION NUMBER: 60/059113  
 ; PRIOR FILING DATE: 1997-09-17  
 ; PRIOR APPLICATION NUMBER: 60/062287  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/063549  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/064103  
 ; PRIOR FILING DATE: 1997-10-31  
 ; PRIOR APPLICATION NUMBER: 60/069873  
 ; PRIOR FILING DATE: 1997-12-17  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/079294  
 ; PRIOR FILING DATE: 1998-03-25  
 ; PRIOR APPLICATION NUMBER: 60/079656  
 ; PRIOR FILING DATE: 1998-03-26  
 ; PRIOR APPLICATION NUMBER: 60/079728  
 ; PRIOR FILING DATE: 1998-03-27  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 246  
 ; SEQ ID NO 155  
 ; LENGTH: 1734  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-232-230-155

Alignment Scores:  
 Pred. No.: 93.3 Length: 1734  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.63% Indels: 0  
 DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-232-230-155 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8  
 DB 1534 ATGGAGCCCCCACACTCCCTCT 1557

RESULT 51

US-10-205-990-201  
 ; Sequence 201, Application US/10205890  
 ; Publication No. US2004004834A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.

```

; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC519
; CURRENT APPLICATION NUMBER: US/10/205,890
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; SEQ ID NO 201
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-205-890-201

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-205-890-201 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8
Db 1534 ATGGAGCCCCCACACTCCCTCCT 1557

RESULT 52
US-10-208-024-201
; Sequence 201, Application US/10208024
; Publication No. US20040048335A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC465
; CURRENT APPLICATION NUMBER: US/10/201,853
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266

```

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC538
; CURRENT APPLICATION NUMBER: US/10/208,024
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; SEQ ID NO 201
; NUMBER OF SEQ ID NOS: 612
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-208-024-201

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-208-024-201 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8
Db 1534 ATGGAGCCCCCACACTCCCTCCT 1557

RESULT 53
US-10-201-853-201
; Sequence 201, Application US/10201853
; Publication No. US20040053358A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC465
; CURRENT APPLICATION NUMBER: US/10/201,853
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266

```



```
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 201
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-201-853-201

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-201-853-201 (1-1734)

Qy 1 MetGlyAlaProThrLeuProPro 8
Db 1534 ATGGGAGCCCCCACACTCCCTCT 1557

RESULT 54
US-10-063-745-51
; Sequence 51, Application US/10063745
; Publication No. US20040059411A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 51
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-745-51

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-063-745-51 (1-1734)
```

```
Qy 1 MetGlyAlaProThrLeuProPro 8
Db 1534 ATGGGAGCCCCCACACTCCCTCT 1557

RESULT 55
US-10-063-512-51
; Sequence 51, Application US/10063512
; Publication No. US20030018183A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 51
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-512-51

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-063-512-51 (1-1734)

Qy 1 MetGlyAlaProThrLeuProPro 8
Db 1534 ATGGGAGCCCCCACACTCCCTCT 1557

RESULT 56
US-10-063-513-51
; Sequence 51, Application US/10063513
; Publication No. US20030018172A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 51
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-513-51
```

Alignment Scores: 93.3 Length: 1734  
Pred. No.: 8.00 Matches: 8  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 5.63% Gaps: 0  
DB: 13

US-09-690-825-34 (1-142) x US-10-063-513-51 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8  
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557

RESULT 57

US-10-063-515-51  
; Sequence 51, Application US/10063515  
; Publication No. US20030018173A1  
; GENERAL INFORMATION:  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3230R1C1  
; CURRENT APPLICATION NUMBER: US/10/063,515  
; CURRENT FILING DATE: 2002-05-01  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 170  
; SEQ ID NO 51  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-063-515-51

Alignment Scores: 93.3 Length: 1734  
Pred. No.: 8.00 Matches: 8  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 5.63% Gaps: 0  
DB: 13

US-09-690-825-34 (1-142) x US-10-063-515-51 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8  
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557

RESULT 58

US-10-063-549-51  
; Sequence 51, Application US/10063549  
; Publication No. US20030027986A1  
; GENERAL INFORMATION:  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3230R1C1  
; CURRENT APPLICATION NUMBER: US/10/063,549

; CURRENT FILING DATE: 2002-05-02  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 170  
; SEQ ID NO 51  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-063-549-51

Alignment Scores: 93.3 Length: 1734  
Pred. No.: 8.00 Matches: 8  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 5.63% Gaps: 0  
DB: 13

US-09-690-825-34 (1-142) x US-10-063-549-51 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8  
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557

RESULT 59

US-10-063-569-51  
; Sequence 51, Application US/10063569  
; Publication No. US20030018168A1  
; GENERAL INFORMATION:  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3230R1C1  
; CURRENT APPLICATION NUMBER: US/10/063,569  
; CURRENT FILING DATE: 2002-05-02  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 170  
; SEQ ID NO 51  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-063-569-51

Alignment Scores: 93.3 Length: 1734  
Pred. No.: 8.00 Matches: 8  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 5.63% Gaps: 0  
DB: 13

US-09-690-825-34 (1-142) x US-10-063-569-51 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8  
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557

RESULT 60

US-10-063-551-51  
; Sequence 51, Application US/10063551  
; Publication No. US20020183494A1  
; GENERAL INFORMATION:  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey

```
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/10/063,551
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 51
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-551-51

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-063-551-51 (1-1734)

Qy 1 MetClyAlaProThrLeuProPro 8
Db 1534 ATGGGAGCCGCCACATCCCTCCT 1557

RESULT 61
US-10-174-581-201
; Sequence 201, Application US/10174581
; Publication No. US20030017540A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C41
; CURRENT APPLICATION NUMBER: US/10/174,581
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/052250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084366
```

; PRIOR FILING DATE: 1998-05-05  
; PRIOR APPLICATION NUMBER: 60/084414  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084639  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084640  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084643  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/085573  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085579  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085580  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085582  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085700  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/086023  
; PRIOR FILING DATE: 1998-05-18  
; PRIOR APPLICATION NUMBER: 60/086392  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/086486  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087098  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087208  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088722  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088740  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088811  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088825  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088863  
; PRIOR FILING DATE: 1998-06-11

; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089090  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653

## Alignment Scores:

Pred. No.:	93.3	Length:	1734
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	13	Gaps:	0

US-09-690-825-34 (1-142) x US-10-174-581-201 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8.

Db 1534 ATGGGAGCCCCCACACTCCCTCT 1557

## RESULT 62

US-10-176-483-201

; Sequence 201, Application US/10176483

; Publication No. US20030017541A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C68

; FILE REFERENCE: P3430R1C68

; CURRENT APPLICATION NUMBER: US/10/176,483

; CURRENT FILING DATE: 2002-06-20

; Prior application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 201

; LENGTH: 1734

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-176-483-201

## Alignment Scores:

Pred. No.:	93.3	Length:	1734
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	13	Gaps:	0

US-09-690-825-34 (1-142) x US-10-176-483-201 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8

Db 1534 ATGGGAGCCCCCACACTCCCTCT 1557

## RESULT 63

US-10-176-749-201  
; Sequence 201, Application US/10176749  
; Publication No. US20030017542A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C76  
; CURRENT APPLICATION NUMBER: US/10/176,749  
; CURRENT FILING DATE: 2002-06-20  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 201  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-176-749-201

## Alignment Scores:

Pred. No.:	93.3	Length:	1734
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	13	Gaps:	0

US-09-690-825-34 (1-142) x US-10-176-749-201 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8  
Db 1534 ATGGAGCCCCCACCTCCCTCT 1557

## RESULT 64

US-10-176-914-201  
; Sequence 201, Application US/10176914  
; Publication No. US20030017543A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C83  
; CURRENT APPLICATION NUMBER: US/10/176,914  
; CURRENT FILING DATE: 2002-06-20  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 201  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-176-914-201

## Alignment Scores:

Pred. No.:	93.3	Length:	1734
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	13	Gaps:	0

US-09-690-825-34 (1-142) x US-10-176-914-201 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8  
Db 1534 ATGGAGCCCCCACCTCCCTCT 1557

## RESULT 65

US-10-176-915-201  
; Sequence 201, Application US/10176915  
; Publication No. US20030017544A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C110  
; CURRENT APPLICATION NUMBER: US/10/176,915  
; CURRENT FILING DATE: 2002-06-21  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 201  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-176-915-201

## Alignment Scores:

Pred. No.:	93.3	Length:	1734
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	13	Gaps:	0

US-09-690-825-34 (1-142) x US-10-176-915-201 (1-1734)

QY 1 MetGlyAlaProThrLeuProPro 8  
Db 1534 ATGGAGCCCCCACCTCCCTCT 1557

## RESULT 66

US-10-006-485A-51  
; Sequence 51, Application US/10006485A  
; Publication No. US20030084062A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan I.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.

APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2830PIC9  
CURRENT APPLICATION NUMBER: US/10/006,485A  
CURRENT FILING DATE: 2001-12-06  
PRIOR APPLICATION NUMBER: 60/098716  
PRIOR FILING DATE: 1998-09-01  
PRIOR APPLICATION NUMBER: 60/098723  
PRIOR FILING DATE: 1998-09-01  
PRIOR APPLICATION NUMBER: 60/098749  
PRIOR FILING DATE: 1998-09-01  
PRIOR APPLICATION NUMBER: 60/098750  
PRIOR FILING DATE: 1998-09-01  
PRIOR APPLICATION NUMBER: 60/098803  
PRIOR FILING DATE: 1998-09-02  
PRIOR APPLICATION NUMBER: 60/098821  
PRIOR FILING DATE: 1998-09-02  
PRIOR APPLICATION NUMBER: 60/098843  
PRIOR FILING DATE: 1998-09-02  
PRIOR APPLICATION NUMBER: 60/098536  
PRIOR FILING DATE: 1998-09-03  
PRIOR APPLICATION NUMBER: 60/098596  
PRIOR FILING DATE: 1998-09-03  
PRIOR APPLICATION NUMBER: 60/098598  
PRIOR FILING DATE: 1998-09-03  
PRIOR APPLICATION NUMBER: 60/098602  
PRIOR FILING DATE: 1998-09-03  
PRIOR APPLICATION NUMBER: 60/098642  
PRIOR FILING DATE: 1998-09-03  
PRIOR APPLICATION NUMBER: 60/098741  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/098754  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/098763  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/098792  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/098808  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/098812  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/098815  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/098816  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/100385  
PRIOR FILING DATE: 1998-09-15  
PRIOR APPLICATION NUMBER: 60/100388  
PRIOR FILING DATE: 1998-09-15  
PRIOR APPLICATION NUMBER: 60/100390  
PRIOR FILING DATE: 1998-09-15  
PRIOR APPLICATION NUMBER: 60/100584  
PRIOR FILING DATE: 1998-09-16  
PRIOR APPLICATION NUMBER: 60/100627  
PRIOR FILING DATE: 1998-09-16  
PRIOR APPLICATION NUMBER: 60/100661  
PRIOR FILING DATE: 1998-09-16  
PRIOR APPLICATION NUMBER: 60/100662  
PRIOR FILING DATE: 1998-09-16  
PRIOR APPLICATION NUMBER: 60/100664  
PRIOR FILING DATE: 1998-09-16  
PRIOR APPLICATION NUMBER: 60/100683  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/100684  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/100710  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/100711  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/100848  
PRIOR FILING DATE: 1998-09-18  
PRIOR APPLICATION NUMBER: 60/100849  
PRIOR FILING DATE: 1998-09-18  
PRIOR APPLICATION NUMBER: 60/100919  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/100930  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/101014  
PRIOR FILING DATE: 1998-09-18  
PRIOR APPLICATION NUMBER: 60/101068  
PRIOR FILING DATE: 1998-09-18  
PRIOR APPLICATION NUMBER: 60/101071  
PRIOR FILING DATE: 1998-09-18  
PRIOR APPLICATION NUMBER: 60/101279  
PRIOR FILING DATE: 1998-09-22  
PRIOR APPLICATION NUMBER: 60/101471  
PRIOR FILING DATE: 1998-09-23  
PRIOR APPLICATION NUMBER: 60/101472  
PRIOR FILING DATE: 1998-09-23  
PRIOR APPLICATION NUMBER: 60/101474  
PRIOR FILING DATE: 1998-09-23  
PRIOR APPLICATION NUMBER: 60/101475  
PRIOR FILING DATE: 1998-09-23  
PRIOR APPLICATION NUMBER: 60/101476  
PRIOR FILING DATE: 1998-09-23  
PRIOR APPLICATION NUMBER: 60/101477  
PRIOR FILING DATE: 1998-09-23  
PRIOR APPLICATION NUMBER: 60/101479  
PRIOR FILING DATE: 1998-09-23  
PRIOR APPLICATION NUMBER: 60/101738  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101741  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101743  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101915  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101916  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/102207  
PRIOR FILING DATE: 1998-09-29  
PRIOR APPLICATION NUMBER: 60/102240  
PRIOR FILING DATE: 1998-09-29  
PRIOR APPLICATION NUMBER: 60/102307  
PRIOR FILING DATE: 1998-09-29  
PRIOR APPLICATION NUMBER: 60/102330  
PRIOR FILING DATE: 1998-09-29  
PRIOR APPLICATION NUMBER: 60/102331  
PRIOR FILING DATE: 1998-09-29  
PRIOR APPLICATION NUMBER: 60/102484  
PRIOR FILING DATE: 1998-09-30  
PRIOR APPLICATION NUMBER: 60/102487  
PRIOR FILING DATE: 1998-09-30  
PRIOR APPLICATION NUMBER: 60/102570  
PRIOR FILING DATE: 1998-09-30  
PRIOR APPLICATION NUMBER: 60/102571  
PRIOR FILING DATE: 1998-09-30  
PRIOR APPLICATION NUMBER: 60/102684  
PRIOR FILING DATE: 1998-10-01  
PRIOR APPLICATION NUMBER: 60/102687  
PRIOR FILING DATE: 1998-10-01  
PRIOR APPLICATION NUMBER: 60/102965  
PRIOR FILING DATE: 1998-10-02  
PRIOR APPLICATION NUMBER: 60/103258  
PRIOR FILING DATE: 1998-10-06  
PRIOR APPLICATION NUMBER: 60/103314  
PRIOR FILING DATE: 1998-10-07  
PRIOR APPLICATION NUMBER: 60/103315  
PRIOR FILING DATE: 1998-10-07  
PRIOR APPLICATION NUMBER: 60/103328  
PRIOR FILING DATE: 1998-10-07  
PRIOR APPLICATION NUMBER: 60/103395  
PRIOR FILING DATE: 1998-10-07  
PRIOR APPLICATION NUMBER: 60/103396

FILE OF INVENTION: 1998-10-07  
FILE OF INVENTION: 60/103401  
FILE OF INVENTION: 1998-10-07  
FILE OF INVENTION: 60/103449  
FILE OF INVENTION: 1998-10-06  
FILE OF INVENTION: 60/103633  
FILE OF INVENTION: 1998-10-08  
FILE OF INVENTION: 60/103678  
FILE OF INVENTION: 1998-10-08  
FILE OF INVENTION: 60/103679  
FILE OF INVENTION: 1998-10-08  
FILE OF INVENTION: 60/103711  
FILE OF INVENTION: 1998-10-08  
FILE OF INVENTION: 60/104257  
FILE OF INVENTION: 1998-10-14  
FILE OF INVENTION: 60/104987  
FILE OF INVENTION: 1998-10-20  
FILE OF INVENTION: 60/105000  
FILE OF INVENTION: 1998-10-20  
FILE OF INVENTION: 60/105002  
FILE OF INVENTION: 1998-10-20  
FILE OF INVENTION: 60/105104  
FILE OF INVENTION: 1998-10-21  
FILE OF INVENTION: 60/105169  
FILE OF INVENTION: 1998-10-22  
FILE OF INVENTION: 60/105266  
FILE OF INVENTION: 1998-10-22  
FILE OF INVENTION: 60/105693  
FILE OF INVENTION: 1998-10-26  
FILE OF INVENTION: 60/105694  
FILE OF INVENTION: 1998-10-26  
FILE OF INVENTION: 60/105807  
FILE OF INVENTION: 1998-10-27  
FILE OF INVENTION: 60/105881  
FILE OF INVENTION: 1998-10-27  
FILE OF INVENTION: 60/105882  
FILE OF INVENTION: 1998-10-27  
FILE OF INVENTION: 60/106023  
FILE OF INVENTION: 1998-10-28

Alignment Scores:  
Pred. No.: 93.3 Length: 1734  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-006-485A-51 (1-1734)

Qy 1 MetGlyAlaProThrLeuProPro 8  
Db 1534 ATGGAGGCCCCACACTCCCTCT 1557

RESULT 67

US-10-013-907A-51  
Sequence 51, Application US/10013907A  
Publication No. US20030064925A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan I.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.

FILE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2830PIC34  
CURRENT APPLICATION NUMBER: US/10/013,907A  
CURRENT FILING DATE: 2001-12-10  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 477  
SEQ ID NO 51  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-013-907A-51  
Alignment Scores:  
Pred. No.: 93.3 Length: 1734  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-013-907A-51 (1-1734)

Qy 1 MetGlyAlaProThrLeuProPro 8  
Db 1534 ATGGAGGCCCCACACTCCCTCT 1557

RESULT 68

US-10-015-499A-51  
Sequence 51, Application US/10015499A  
Publication No. US20030065142A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan I.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.

FILE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2830PIC42  
CURRENT APPLICATION NUMBER: US/10/015,499A  
CURRENT FILING DATE: 2001-12-11  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 477  
SEQ ID NO 51  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-015-499A-51

Alignment Scores:  
Pred. No.: 93.3 Length: 1734  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-015-499A-51 (1-1734)

Qy 1 MetGlyAlaProThrLeuProPro 8  
Db 1534 ATGGAGGCCCCACACTCCCTCT 1557

```

RESULT 69
US-10-063-555-51
; Sequence 51, Application US/10063555
; Publication No. US20030065143A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,555
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 51
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-555-51

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-063-555-51 (1-1734)
QY 1 MetGlyAlaProThrLeuPro 8
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557

RESULT 70
US-10-063-563-51
; Sequence 51, Application US/10063563
; Publication No. US20030060602A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,563
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 51
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-563-51

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-063-563-51 (1-1734)
QY 1 MetGlyAlaProThrLeuPro 8
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557

```

```

Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-063-563-51 (1-1734)
QY 1 MetGlyAlaProThrLeuPro 8
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557

RESULT 71
US-10-063-594-51
; Sequence 51, Application US/10063594
; Publication No. US20030065161A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,594
; CURRENT FILING DATE: 2002-05-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 51
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-594-51

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-063-594-51 (1-1734)
QY 1 MetGlyAlaProThrLeuPro 8
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557

RESULT 72
US-10-063-553-51
; Sequence 51, Application US/10063553
; Publication No. US20030045684A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,553
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 51
; LENGTH: 1734

```



```

; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-553-51

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-063-553-51 (1-1734)
QY 1 MetGlyAlaProThrLeuProPro 8
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557

RESULT 73
US-10-063-554-51
; Sequence 51, Application US/10063554
; Publication No. US20030040013A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,554
; CURRENT FILING DATE: 2002-05-02
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 51
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-554-51

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-063-554-51 (1-1734)
QY 1 MetGlyAlaProThrLeuProPro 8
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557

RESULT 74
US-10-176-484-201
; Sequence 201, Application US/10176484
; Publication No. US20030059876A9
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C149
; CURRENT APPLICATION NUMBER: US/10/180,550
; CURRENT FILING DATE: 2002-06-25
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 201
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-180-550-201

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-180-550-201 (1-1734)
QY 1 MetGlyAlaProThrLeuProPro 8
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557

RESULT 75
US-10-180-550-201
; Sequence 201, Application US/10180550
; Publication No. US2003006440A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C149
; CURRENT APPLICATION NUMBER: US/10/180,550
; CURRENT FILING DATE: 2002-06-25
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 201
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-180-550-201

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-176-484-201 (1-1734)
QY 1 MetGlyAlaProThrLeuProPro 8
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557

RESULT 76
US-10-176-484-201
; Sequence 201, Application US/10176484
; Publication No. US20030059876A9
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C149
; CURRENT APPLICATION NUMBER: US/10/180,550
; CURRENT FILING DATE: 2002-06-25
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 201
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-180-550-201

Alignment Scores:
Pred. No.: 93.3 Length: 1734
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 13 Gaps: 0

US-09-690-825-34 (1-142) x US-10-180-550-201 (1-1734)
QY 1 MetGlyAlaProThrLeuProPro 8
Db 1534 ATGGGAGCCCCCACACTCCCTCCT 1557
```

Tue Aug 17 05:54:52 2004

us-09-690-825-34.oligo.rnpb

Page 32

Search completed: August 16, 2004, 05:21:35  
Job time : 1494 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 16, 2004, 01:19:04 ; Search time 96 Seconds  
(without alignments)  
820.865 Million cell updates/sec

Title: US-09-690-825-34

Perfect score: 142  
Sequence: 1 MGRPTLPANQPIKOHRS.....EPEETAKVRRALQLAAMD 142

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1360442

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09690825/runat\_11082004\_141015\_13818/app\_query.fasta\_1.327  
-DB=Issued Patents NA -QPM=fastap -SUFFIX=ologo.rn1 -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=100 -DOCLALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL  
-OUTFMT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09690825 -CGN 1\_1\_69 @runat\_11082004\_141015\_13818 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142	100.0	426	4	US-09-283-144-2
2	142	100.0	1619	3	US-09-163-162-1
3	142	100.0	1619	3	US-09-286-407-1
4	142	100.0	1619	4	US-09-496-694B-97
C 5	107	75.4	1165	3	US-08-448-722A-1
C 6	107	75.4	1165	3	US-08-189-309B-1
7	93	55.5	740	4	US-09-016-434-918
8	41	28.9	955	4	US-09-496-694B-10
9	39	27.5	14796	3	US-08-975-080-35
10	39	27.5	14796	3	US-09-630-706-10
11	39	27.5	14796	4	US-09-496-694B-3
12	37	26.1	417	4	US-09-283-144-1

13	14	9.9	332	4	US-09-833-381-1331	Sequence 1331, Ap
14	9	6.3	30	4	US-09-672-717-199	Sequence 199, App
15	8	5.6	26	4	US-09-496-694B-13	Sequence 13, Appl
16	8	5.6	281	4	US-09-833-381-1330	Sequence 1330, Ap
17	8	5.6	293	4	US-09-621-976-13359	Sequence 13359, A
18	8	5.6	326	4	US-09-621-976-13398	Sequence 13398, A
19	8	5.6	337	4	US-09-621-976-13384	Sequence 13384, A
20	8	5.6	342	4	US-09-621-976-13367	Sequence 13367, A
21	8	5.6	346	4	US-09-621-976-13443	Sequence 13443, A
22	8	5.6	349	4	US-09-621-976-13375	Sequence 13375, A
23	8	5.6	351	4	US-09-621-976-13420	Sequence 13420, A
24	8	5.6	357	4	US-09-621-976-13407	Sequence 13407, A
25	8	5.6	361	4	US-09-621-976-13392	Sequence 13392, A
26	8	5.6	365	4	US-09-621-976-13428	Sequence 13428, A
27	8	5.6	515	4	US-09-496-694B-211	Sequence 231, App
28	8	5.6	4376	1	US-08-119-125A-1	Sequence 1, Appl
29	8	5.6	6744	1	US-08-119-125A-2	Sequence 2, Appl
30	7	4.9	21	3	US-09-163-162-2	Sequence 2, Appl
31	7	4.9	21	3	US-09-286-407-2	Sequence 2, Appl
32	7	4.9	21	4	US-09-496-694B-4	Sequence 4, Appl
C 33	7	4.9	22	3	US-09-163-162-3	Sequence 3, Appl
C 34	7	4.9	22	3	US-09-286-407-3	Sequence 3, Appl
C 35	7	4.9	22	4	US-09-496-694B-5	Sequence 5, Appl
36	7	4.9	23	3	US-09-163-162-4	Sequence 4, Appl
37	7	4.9	23	3	US-09-286-407-4	Sequence 4, Appl
38	7	4.9	23	4	US-09-496-694B-6	Sequence 6, Appl
39	7	4.9	273	4	US-09-313-294A-1818	Sequence 1818, Ap
C 40	7	4.9	310	3	US-09-328-111-715	Sequence 715, App
41	7	4.9	504	4	US-09-107-532A-1678	Sequence 1678, Ap
42	7	4.9	559	4	US-09-621-976-10449	Sequence 10449, A
43	7	4.9	561	4	US-09-107-532A-1373	Sequence 1373, Ap
44	7	4.9	588	4	US-09-252-991A-4396	Sequence 4396, Ap
C 45	7	4.9	611	3	US-08-642-274D-16	Sequence 16, Appl
C 46	7	4.9	611	3	US-08-952-014C-16	Sequence 16, Appl
47	7	4.9	627	4	US-09-134-000C-912	Sequence 912, App
C 48	7	4.9	642	4	US-09-489-039A-6367	Sequence 6367, Ap
49	7	4.9	777	4	US-09-134-000C-375	Sequence 375, App
C 50	7	4.9	924	4	US-09-252-991A-4263	Sequence 4263, Ap
C 51	7	4.9	939	4	US-09-252-991A-3424	Sequence 3424, Ap
52	7	4.9	942	3	US-09-475-316A-117	Sequence 117, App
53	7	4.9	942	4	US-09-704-640-117	Sequence 117, App
54	7	4.9	944	3	US-09-475-316A-119	Sequence 119, App
55	7	4.9	944	4	US-09-704-640-119	Sequence 119, App
56	7	4.9	951	4	US-09-489-039A-5085	Sequence 5085, Ap
C 57	7	4.9	993	3	US-08-969-644-21	Sequence 21, Appl
C 58	7	4.9	993	3	US-08-444-189-21	Sequence 21, Appl
C 59	7	4.9	993	3	US-08-468-544-21	Sequence 21, Appl
60	7	4.9	1095	2	US-08-578-404B-9	Sequence 9, Appl
61	7	4.9	1190	3	US-09-475-316A-61	Sequence 61, Appl
62	7	4.9	1190	4	US-09-704-640-61	Sequence 61, Appl
C 63	7	4.9	1194	4	US-09-328-352-3331	Sequence 3331, Ap
C 64	7	4.9	1230	4	US-09-252-991A-8373	Sequence 8373, Ap
65	7	4.9	1257	4	US-09-252-991A-8504	Sequence 8504, Ap
66	7	4.9	1282	3	US-09-475-316A-69	Sequence 69, Appl
67	7	4.9	1282	4	US-09-704-640-69	Sequence 69, Appl
68	7	4.9	1308	3	US-09-475-316A-65	Sequence 65, Appl
69	7	4.9	1308	4	US-09-704-640-65	Sequence 65, Appl
70	7	4.9	1328	3	US-09-475-316A-71	Sequence 71, Appl
71	7	4.9	1328	4	US-09-704-640-71	Sequence 71, Appl
72	7	4.9	1551	4	US-09-205-258-212	Sequence 212, App
73	7	4.9	1644	4	US-09-548-004-9	Sequence 9, Appl
74	7	4.9	1743	4	US-09-107-532A-2080	Sequence 2080, Ap
75	7	4.9	1779	1	US-08-208-036-16	Sequence 16, Appl
76	7	4.9	1779	1	US-08-428-823-16	Sequence 16, Appl
77	7	4.9	2463	4	US-09-134-000C-3032	Sequence 3032, Ap
C 78	7	4.9	2517	1	US-07-906-930B-1	Sequence 1, Appl
79	7	4.9	2631	1	US-08-428-036-13	Sequence 13, Appl
80	7	4.9	2631	1	US-08-428-823-13	Sequence 13, Appl
C 81	7	4.9	2939	1	US-07-906-930B-3	Sequence 3, Appl
82	7	4.9	2997	4	US-09-252-991A-4487	Sequence 4487, Ap
83	7	4.9	3252	1	US-08-208-036-11	Sequence 11, Appl
84	7	4.9	3252	1	US-08-428-823-11	Sequence 11, Appl
C 85	7	4.9	3465	4	US-09-023-655-1136	Sequence 1136, Ap

86	7	4.9	4792	4	US-08-361-527-156	Sequence 156, Appl
87	7	4.9	4934	4	US-09-321-017B-985	Sequence 985, Appl
88	7	4.9	6835	4	US-09-125-633-1	Sequence 1, Appl
89	7	4.9	7100	4	US-09-308-375-1	Sequence 1, Appl
90	7	4.9	7502	3	US-08-969-644-6	Sequence 6, Appl
91	7	4.9	7502	3	US-08-444-189-6	Sequence 6, Appl
92	7	4.9	7502	3	US-08-468-544-6	Sequence 6, Appl
93	7	4.9	9361	4	US-08-934-386-7	Sequence 6, Appl
94	7	4.9	17417	4	US-09-648-004-27	Sequence 27, Appl
95	7	4.9	30001	1	US-08-125-468-1	Sequence 1, Appl
96	7	4.9	30001	2	US-08-474-933-1	Sequence 1, Appl
97	7	4.9	31571	1	US-08-323-443B-1	Sequence 1, Appl
98	7	4.9	53526	3	US-08-658-136-2	Sequence 2, Appl
99	7	4.9	53577	3	US-08-658-136-1	Sequence 1, Appl
100	7	4.9	152331	3	US-09-128-153-15	Sequence 16, Appl

## ALIGNMENTS

```

RESULT 1
US-09-283-144-2
; Sequence 2, Application US/09283144
; Patent No. 6346389
; GENERAL INFORMATION:
; APPLICANT: Yale University
; TITLE OF INVENTION: Method for Selectively Modulating the Interactions
; TITLE OF INVENTION: between Survivin and Tubulin
; FILE REFERENCE: 44574-5033-US
; CURRENT APPLICATION NUMBER: US/09/283,144
; CURRENT FILING DATE: 1999-04-01
; EARLIER APPLICATION NUMBER: US 60/080,288
; EARLIER FILING DATE: 1998-04-01
; EARLIER APPLICATION NUMBER: US 08/975,080
; EARLIER FILING DATE: 1997-11-20
; EARLIER APPLICATION NUMBER: PCT/US97/21880
; EARLIER FILING DATE: 1997-11-20
; EARLIER APPLICATION NUMBER: US 60/031,435
; EARLIER FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of open reading frame encoding
; US-09-283-144-2

```

```

Db      181  TTCAAGGAGCTGGAAGGCTGGGAGCCAGATCAGCACCCCATAGAGGAACATAAAAAAGCAT 240
Qy      81  SerSerGlyCysAlapheLeuSerVallyslYsGlnPheGluGluLeuThrLeuGlyGlu 100
Db      241  TCGTCGGTGTGGCTTTCTCTTGTCAAGAGCAGTTTGAAGAAATTAACCCCTTGGTGAA 300
Qy      101  PheLeuLeuLeuAspArgGluArgAlaLysAsnLysileAlaLysGluThrAsnAsnLys 120
Db      301  TTTTTCGAACTGCAGACAGAAAGCCCAAGAACAAAATTCCAAAGGAAACCAACATAAG 360
Qy      121  LysLysGluPheGluGluThrAlaLysLysValArgArgAlaIleGluGlnLeuAla 140
Db      361  AGAAGAAGAAATTTGAGGAAACTGCGAAGAAAGTGGCGCGTCCCATCGAGCAGCTGGCTGCC 420
Qy      141  MetAsp 142
Db      421  ATGGAT 426

RESULT 2
US-09-163-162-1
; Sequence 1, Application US/09163162
; Patent No. 607789
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Ackermann, Elizabeth J.
; APPLICANT: Swayze, Eric E.
; APPLICANT: Cowsett, Lex M.
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: RTS-0008
; CURRENT APPLICATION NUMBER: US/09/163.162
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)..(478)
US-09-163-162-1

Alignment Scores:
Pred. No.: 4.84e-140
Score: 142.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 3
US-09-690-825-34 (1-142) x US-09-163-162-1 (1-1619)

```

Db 350 TTTTGAAGTGGACAGAGAAAGACCCAGAACAAAATTGCAAGAGAAACCAACAATAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGluLeuAlaAla 140  
Db 410 AAGAAAGAAATTTGAGAAACTGCGAAGAAAGTGGCCCGTGCAATCGACAGCTGGCTGCC 469  
QY 141 MetAsp 142  
Db 470 ATGGAT 475  
RESULT 3  
US-09-286-407-1  
; Sequence 1, Application US/09286407A  
; Patent No. 6165788  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, C. Frank  
; APPLICANT: Ackermann, Elizabeth J.  
; APPLICANT: Swayze, Eric E.  
; APPLICANT: Cowsett, Lex M.  
; TITLE OF INVENTION: ANTISENSE MODULATION OF Survivin EXPRESSION  
; FILE REFERENCE: ISPH-0349  
; CURRENT APPLICATION NUMBER: US/09/286,407A  
; CURRENT FILING DATE: 1999-04-05  
; NUMBER OF SEQ ID NOS: 48  
; SEQ ID NO 1  
; LENGTH: 1619  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (50)...(478)  
US-09-286-407-1  
Alignment Scores:  
Pred. No.: 4,84e-140 Length: 1619  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 3  
DB: 1  
US-09-690-825-34 (1-142) x US-09-286-407-1 (1-1619)  
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 50 ATGGGTGCCCGACGTTGCCCTGCTGGCAGCCCTTTCTCAAGACACCCGATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 110 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGGCCCTGCACCCCGAGCGGATGCCCGAG 169  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
Db 170 GCTGGCTTCATCCACTGCCCTGAGACGAGCCAGACTTGGCCAGTGTCTTCTGTC 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspPheProIleGluHisLysLysHis 80  
Db 230 TTCAGGAGCTGGAGGCTGGAGCCAGATGACGCCCATAGAGGAACATAAAAGCAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
Db 290 TCGTCCGGTTCGGCTTTCCTTCTGTCAAGAGGAGTTTGAAGAAATTAACCTTGGTGAA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
Db 350 TTTTGAACCTGCACAGAGAAAGGCCAAGAACAAAATTGCAAGGAAACCAACAATAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGluLeuAlaAla 140  
Db 410 AAGAAAGAAATTTGAGGAAACTGCGCAAGAAAGTGGCCCGTGCCATCGAGCAGCTGGCTGCC 469  
QY 141 MetAsp 142  
Db 470 ATGGAT 475

Db 470 ATGGAT 475  
RESULT 4  
US-09-496-694B-97  
; Sequence 97, Application US/09496694B  
; Patent No. 6335194  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Elizabeth J. Ackermann  
; APPLICANT: Eric E. Swayze  
; APPLICANT: Lex M. Cowsett  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
; FILE REFERENCE: ISPH-0439  
; CURRENT APPLICATION NUMBER: US/09/496,694B  
; CURRENT FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 09/286,407  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 09/163,162  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 249  
; SEQ ID NO 97  
; LENGTH: 1619  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-496-694B-97  
Alignment Scores:  
Pred. No.: 4,84e-140 Length: 1619  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 4  
DB: 4  
US-09-690-825-34 (1-142) x US-09-496-694B-97 (1-1619)  
QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 50 ATGGGTGCCCGACGTTGCCCTGCTGGCAGCCCTTTCTCAAGGACACCCGATCTCT 109  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 110 ACATTCAAGAACTGGCCCTTCTTGGAGGGCTGGCCCTGCACCCCGAGCGGATGCCCGAG 169  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPheCys 60  
Db 170 GCTGGCTTCATCCACTGCCCTGAGACGAGCCAGACTTGGCCAGTGTCTTCTGTC 229  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspPheProIleGluHisLysLysHis 80  
Db 230 TTCAGGAGCTGGAGGCTGGAGCCAGATGACGCCCATAGAGGAACATAAAAGCAT 289  
QY 81 SerSerGlyCysAlaPheLeuSerValLysLysGlnPheGluGluLeuThrLeuGlyGlu 100  
Db 290 TCGTCCGGTTCGGCTTTCCTTCTGTCAAGAGGAGTTTGAAGAAATTAACCTTGGTGAA 349  
QY 101 PheLeuLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnLys 120  
Db 350 TTTTGAACCTGCACAGAGAAAGGCCAAGAACAAAATTGCAAGGAAACCAACAATAAG 409  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGluLeuAlaAla 140  
Db 410 AAGAAAGAAATTTGAGGAAACTGCGCAAGAAAGTGGCCCGTGCCATCGAGCAGCTGGCTGCC 469  
QY 141 MetAsp 142  
Db 470 ATGGAT 475  
RESULT 5  
US-09-448-722A-1/c  
; Sequence 1, Application US/08448722A  
; Patent No. 6072028  
; GENERAL INFORMATION:



QY 1 MetGlyAlaProThrLeuProAlaTrpGlnProPheLeuLysAspHisArgIleSer 20  
Db 1114 ATGGGTGCCCCGACGTTGCCCCCTGCCCTGGCAGCCCTTCTCAAGGACCCACCGCATCTCT 1055  
QY 21 ThrPheLysAsnTrpProPheLeuGluGlyCysAlaCysThrProGluArgMetAlaGlu 40  
Db 1054 ACATTCAAGAACTGCCCTTCTTGTGAGGGCTGGCG-TGCACCCCGAGCGGATGGCCGAG 996  
QY 41 AlaGlyPheIleHisCysProThrGluAsnGluProAspLeuAlaGlnCysPhePheCys 60  
Db 995 GCTGGCTTCATCCACTGCCCTCAAGAACGAGCAGACTTGGCCCGAGTGTTCCTTCGCG 936  
QY 61 PheLysGluLeuGluGlyTrpGluProAspAspProIleGluGluHisIleLysHis 80  
Db 935 TTCAGGAGCTGGAAGGCTGGAGCGCAGATGACGCCACCCATAGAGAAACATAAAAGCAT 876  
QY 81 SerSerGlyCysAlaPheLeuSerValIleLysGlnPheGluGluLeuThrIleuGlyGlu 100  
Db 875 TCGTCGGTGGCTTCTCTCTGTCAGAGAGCAGTTTGAAGAATTAAACCTTGGTGAA 816  
QY 101 PheLysLeuAspArgGluArgAlaLysAsnLysIleAlaLysGluThrAsnAsnLys 120  
Db 815 TTTTGAAGAACTGGACAGAGAAAGCCCAAGAACAAATTGCAAGAAACCCACACANTAG 756  
QY 121 LysLysGluPheGluGluThrAlaLysLysValArgAlaIleGluGluLeuAla 139  
Db 755 AAGAAAGAAATTTCAGGAAACTGGAAGAAAGTGGCGCGTGCATCGCAGCGTGGCC 699

RESULT 7  
US-09-016-434-918  
; Sequence 918, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,434  
; FILING DATE: HEREMITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0002 US  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 918:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 740 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: BRAITUT01

CLONE: 752848  
US-09-016-434-918  
Alignment Scores:  
Pred. No.: 7,26e-89 Length: 740  
Score: 93.00 Matches: 130  
Percent Similarity: 98.48% Conservative: 0  
Best Local Similarity: 98.48% Mismatches: 1  
Query Match: 65.49% Indels: 2  
DB: 4 Gaps: 0  
US-09-690-825-34 (1-142) x US-09-016-434-918 (1-740)  
QY 12 ProPheLeuLysAspHisArgIleSerThrPheLysAsnTrpProPheLeuGluGlyCys 31  
Db 51 CCCCCCTCAAGGACCCACCGCATCTCTACATTCAAGAACTGGCCCTTCTTGGAGGCTGC 110  
QY 32 AlaCysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrGlu-AsnGlu 51  
Db 111 GCTGTCACCCCGAGCGGATGGCGGAGCTGGCTTATCCACTGCCCTCTGA-NACGA 169  
QY 51 uProAspLeuAlaGlnCysPhePheCysPheLysGluLeuGluGlyTrpGluProAspAs 71  
Db 170 GCCAGACTTGGCCAGTGTTCCTCTGCTTCAAGGAGCTGGAAGGCTGGAGCCAGATGA 229  
QY 71 pAspProIleGluGluHisLysLysHisSerSerGlyCysAlaPheLeuSerValIleLys 91  
Db 230 CGACCCCATAGAGAACATAAAAGCATTCGTCGGTGGCTTCTTCTGTCAAGAA 289  
QY 91 sGlnPheGluGluLeuThrLeuGlyGluPheLeuLysLeuAspArgGluArgAlaLysAs 111  
Db 290 GCATTTGAAGAAATTAAACCTTGGTGAATTTTGAAGAACTGCACAGAGAAAGCCAAAGAA 349  
QY 111 nLysIleAlaLysGluThrAsnAsnLysLysLysGluPheGluGluThrAlaLysLysVa 131  
Db 350 CAAATTTGCAAGGAAACCAACAAATAAGAAAGAAATTTGAGGAAACTCGGAAGAAAGT 409  
QY 131 lArgArgAlaIleGluGluLeuAlaAlaMetAsp 142  
Db 410 GCGCGTGCATCGAGCAGCTGGCTGCCATGGAT 443

RESULT 8  
US-09-496-694B-10  
; Sequence 10, Application US/09496694B  
; Patent No. 6335194  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Elizabeth J. Ackermann  
; APPLICANT: Eric E. Swazy  
; APPLICANT: Lex M. Cowser  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
; FILE REFERENCE: ISPH-0439  
; CURRENT APPLICATION NUMBER: US/09/496,694B  
; CURRENT FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 09/286,407  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 09/163,162  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 249  
; SEQ ID NO 10  
; LENGTH: 955  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (109)...(531)  
US-09-496-694B-10  
Alignment Scores:  
Pred. No.: 4.04e-34 Length: 955  
Score: 41.00 Matches: 41  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

QY	DB	US-09-690-825-34 (1-142) x US-09-496-694B-10 (1-955)	Query Match:	28.87%	Indels:	0
			DB:	4	Gaps:	0
QY	31	CysAlaCysThrProGluArgMetAlaGluAlaGlyPheIleHisCysProThrGluAsn 50				
Db	199	TGGCGCTGCACCCAGAGGAATGCGGAGGTGCGCTTCATCCACTGCCCTACCGAAGC 258				
QY	51	GluProAspLeuAlaGlnCysPheCysPheLeuGluGlyTrpGluProAsp 70				
Db	259	GAGCCTGATTGGCCAGTGTTTTCTGCTTTAAGGATTTGAAGGCTGGGAACCGAT 318				
QY	71	Asp 71				
Db	319	GAC 321				
<p>RESULT 9</p> <p>US-08-975-080-35</p> <p>; Sequence 35, Application US/08975080</p> <p>; Patent No. 6245523</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Altieri, Dario C.</p> <p>; TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS</p> <p>; TITLE OF INVENTION: CELLULAR APOPTOSIS, AND ITS MODULATION</p> <p>; NUMBER OF SEQUENCES: 35</p> <p>; CORRESPONDENCE ADDRESS:</p> <p>; ADDRESSEE: MORGAN, LEWIS &amp; BOCKIUS LLP</p> <p>; STREET: 1800 M Street, N.W.</p> <p>; CITY: Washington</p> <p>; STATE: D.C.</p> <p>; COUNTRY: USA</p> <p>; ZIP: 20036-5869</p> <p>; COMPUTER READABLE FORM:</p> <p>; MEDIUM TYPE: Floppy disk</p> <p>; COMPUTER: IBM PC compatible</p> <p>; OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>; SOFTWARE: Patent In Release #1.0, Version #1.30</p> <p>; CURRENT APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US/08/975,080</p> <p>; FILING DATE: 20-NOV-1997</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US 60/031,435</p> <p>; FILING DATE: 20-NOV-1996</p> <p>; ATTORNEY/AGENT INFORMATION:</p> <p>; NAME: Adler, Reid G.</p> <p>; REGISTRATION NUMBER: 30,988</p> <p>; REFERENCE/DOCKET NUMBER: 044574-5022-01-WO</p> <p>; TELECOMMUNICATION INFORMATION:</p> <p>; TELEPHONE: 202-467-7000</p> <p>; TELEFAX: 202-467-7176</p> <p>; INFORMATION FOR SEQ ID NO: 35:</p> <p>; SEQUENCE CHARACTERISTICS:</p> <p>; LENGTH: 14796 base pairs</p> <p>; TYPE: nucleic acid</p> <p>; STRANDEDNESS: single</p> <p>; TOPOLOGY: linear</p> <p>; MOLECULE TYPE: DNA (genomic)</p> <p>US-08-975-080-35</p> <p>Alignment Scores:</p> <p>Pred. No.: 6.48e-31</p> <p>Score: 14796</p> <p>Percent Similarity: 39.00%</p> <p>Best Local Similarity: 100.00%</p> <p>Query Match: 27.46%</p> <p>DB: 3</p>						
QY	75	GluGluHisLysLysHisSerGlyCysAlaPheLeuSerValLysLysGlnPheGlu 94				
Db	5159	GAGGAACATAAAAGCATTGTCGCTTTCCTTCTGTCGAAGAGCAGTTTCAA 5218				
QY	95	GluLeuThrLeuGlyGluPheLeuLysLeuAspArgGluArgAlaLysAsnLysIle 113				
Db	5219	GAATTAACCTTGGTGAATTTTGAAGAACTGGACAGAGAAGAGCCAGCAAAATT 5275				
<p>RESULT 11</p> <p>US-09-496-694B-3</p> <p>; Sequence 3, Application US/09496694B</p> <p>; Patent No. 6335194</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: C. Frank Bennett</p> <p>; APPLICANT: Elizabeth J. Ackermann</p> <p>; APPLICANT: Eric E. Swayze</p> <p>; APPLICANT: Lex M. Cowsett</p> <p>; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION</p> <p>; FILE REFERENCE: ISPH-0439</p> <p>; CURRENT APPLICATION NUMBER: US/09/496,694B</p> <p>; CURRENT FILING DATE: 2000-02-02</p> <p>; PRIOR APPLICATION NUMBER: 09/286,407</p> <p>; PRIOR FILING DATE: 1999-04-05</p> <p>; PRIOR APPLICATION NUMBER: 09/163,162</p> <p>; PRIOR FILING DATE: 1998-09-29</p> <p>; NUMBER OF SEQ ID NOS: 249</p> <p>; SEQ ID NO 3</p> <p>; LENGTH: 14796</p> <p>; TYPE: DNA</p> <p>; ORGANISM: Homo sapiens</p> <p>; FEATURE:</p>						



/	NAME/KEY:	CDS
/	LOCATION:	(2811)...(2921)
/	NAME/KEY:	CDS
/	LOCATION:	(3174)...(3283)
/	NAME/KEY:	CDS
/	LOCATION:	(5158)...(5275)
/	NAME/KEY:	CDS
/	LOCATION:	(11955)...(12044)
/	US-09-496-694B-3	
Alignment Scores:		
Pred. No.:	6,48e-31	Length: 14796
Score:	39.00	Matches: 39
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	27.46%	Indels: 0
DB:	4	Gaps: 0
US-09-690-825-34 (1-142) x US-09-496-694B-3 (1-14796)		
QY	75	GlucHuhisLysHisSerSergLyCyAAlAPheLeuSerVallyLysGlnPheGlu 94
Db	5159	GAGGAACATAAAAGCATTCGCGGTGGCTTCCCTTTCTGCAAGAAGCAGTTTGAA 5218
QY	95	GlueuThrLeuGlyGluPheLeuLysLeuAspArgGluArgAlaLysAsnLysIle 113
Db	5219	GAATTAACCTTGTTGTAATTITGAACACTGGACAGAGAGGCCAAGACAAAATT 5275
RESULT 12		
US-09-283-144-1		
/	Sequence 1,	Application US/09283144
/	Patent No.	6346389
/	GENERAL INFORMATION:	
/	APPLICANT:	Yale University
/	TITLE OF INVENTION:	Method for Selectively Modulating the Interactions
/	TITLE OF INVENTION:	between Survivin and Tubulin
/	FILE REFERENCE:	44574-5033-US
/	CURRENT APPLICATION NUMBER:	US/09/283,144
/	CURRENT FILING DATE:	1999-04-01
/	EARLIER APPLICATION NUMBER:	US 60/080,288
/	EARLIER FILING DATE:	1998-04-01
/	EARLIER APPLICATION NUMBER:	US 08/975,080
/	EARLIER FILING DATE:	1997-11-20
/	EARLIER APPLICATION NUMBER:	PCT/US97/21880
/	EARLIER FILING DATE:	1997-11-20
/	EARLIER APPLICATION NUMBER:	US 60/031,435
/	EARLIER FILING DATE:	1996-11-20
/	NUMBER OF SEQ ID NOS:	15
/	SOFTWARE:	PatentIn Ver. 2.0
/	SEQ ID NO 1	
/	LENGTH:	417
/	TYPE:	DNA
/	ORGANISM:	Homo sapiens
/	FEATURE:	
/	OTHER INFORMATION: DNA sequence of proximal 5' flanking region of	
/	OTHER INFORMATION: Survivin gene	
/	FEATURE:	
/	NAME/KEY:	CDS
/	LOCATION:	(291)..(401)
/	US-09-283-144-1	
Alignment Scores:		
Pred. No.:	3e-30	Length: 417
Score:	37.00	Matches: 37
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	26.06%	Indels: 0
DB:	4	Gaps: 0
US-09-690-825-34 (1-142) x US-09-283-144-1 (1-417)		
QY	1	MetGlyAlaProThrLeuProProlatrpGlnPropHeLysAspHisArgIleSer 20

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.34% Indels: 0  
 DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-672-717-199 (1-30)

QY 69 ProAspAspProIleGluHis 77

DB 3 CCAGTATGACCCCATAGAGACAT 29

# RESULT 15

US-09-496-694B-13

; Sequence 13, Application US/09496694B

; Patent No. 6335194

; GENERAL INFORMATION:

; APPLICANT: C. Frank Bennett

; APPLICANT: Elizabeth J. Ackermann

; APPLICANT: Eric E. Swayze

; APPLICANT: Lex M. Cowsett

; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION

; FILE REFERENCE: ISPH-0439

; CURRENT APPLICATION NUMBER: US/09/496.694B

; CURRENT FILING DATE: 2000-02-02

; PRIOR APPLICATION NUMBER: 09/286,407

; PRIOR FILING DATE: 1999-04-05

; PRIOR APPLICATION NUMBER: 09/163,162

; PRIOR FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 249

; SEQ ID NO 13

; LENGTH: 26

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: PCR Probe

US-09-496-694B-13

Alignment Scores: 0.682 Length: 26

Pred. No.: 8 Matches: 8

Score: 8.00 Conservat: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 5.63% Indels: 0

DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-496-694B-13 (1-26)

QY 62 LysGluLeuGluGlyTrpGluPro 69

DB 2 AAGGAATTGGAAGGCTGGAAACC 25

# RESULT 16

US-09-833-381-1330

; Sequence 1330, Application US/09833381

; Patent No. 6672186

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 2050

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1330

; LENGTH: 281

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)-(281)

; OTHER INFORMATION: n = A,T,C or G

US-09-833-381-1330

Alignment Scores: 7.1 Length: 326

Pred. No.: 8 Matches: 8

Score: 8.00 Conservat: 0

Best Local Similarity: 100.00%

Alignment Scores: 6.19 Length: 281

Pred. No.: 8 Matches: 8

Score: 8.00 Conservat: 0

Best Local Similarity: 100.00%

Query Match: 5.63% Indels: 0

DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-833-381-1330 (1-281)

QY 60 CysPheLysGluLeuGluGlyTrp 67

DB 258 TCCTTCAGGAGCTGGAGGCTGG 281

# RESULT 17

US-09-621-976-13359

; Sequence 13359, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 13359

; LENGTH: 293

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-621-976-13359

Alignment Scores: 6.44 Length: 293

Pred. No.: 8 Matches: 8

Score: 8.00 Conservat: 0

Best Local Similarity: 100.00%

Query Match: 5.63% Indels: 0

DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-621-976-13359 (1-293)

QY 1 MetGlyAlaProThrLeuProPro 8

DB 237 ATGGGAGCCCCACACCTCTCTCT 260

# RESULT 18

US-09-621-976-13398

; Sequence 13398, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 13398

; LENGTH: 326

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-621-976-13398

Alignment Scores: 7.1 Length: 326

Pred. No.: 8 Matches: 8

Score: 8.00 Conservat: 0

Best Local Similarity: 100.00%

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-621-976-13398 (1-326)

QY 1 MetGlyAlaProThrLeuProPro 8  
DB 237 ATGGAGCCCCCACACTCCCTCCT 260

RESULT 19  
US-09-621-976-13384  
; Sequence 13384, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 13384  
; LENGTH: 337  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-13384

Alignment Scores:  
Pred. No.: 7.33 Length: 337  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-621-976-13384 (1-337)

QY 1 MetGlyAlaProThrLeuProPro 8  
DB 237 ATGGAGCCCCCACACTCCCTCCT 260

RESULT 20  
US-09-621-976-13367  
; Sequence 13367, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 13367  
; LENGTH: 342  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 229  
; OTHER INFORMATION: n=a, g, c or t  
US-09-621-976-13367

Alignment Scores:  
Pred. No.: 7.43 Length: 342  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 5.63% Indels: 0  
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-621-976-13367 (1-342)

QY 1 MetGlyAlaProThrLeuProPro 8  
DB 237 ATGGAGCCCCCACACTCCCTCCT 260

RESULT 21  
US-09-621-976-13443  
; Sequence 13443, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 13443  
; LENGTH: 346  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 254..255  
; OTHER INFORMATION: n=a, g, c or t  
US-09-621-976-13443

Alignment Scores:  
Pred. No.: 7.51 Length: 346  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-621-976-13443 (1-346)

QY 1 MetGlyAlaProThrLeuProPro 8  
DB 262 ATGGAGCCCCCACACTCCCTCCT 285

RESULT 22  
US-09-621-976-13375  
; Sequence 13375, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 13375  
; LENGTH: 349  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-13375

Alignment Scores:  
Pred. No.: 7.57 Length: 349  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0

```
DB: 4 Gaps: 0
US-09-690-825-34 (1-142) x US-09-621-976-13375 (1-349)
QY 1 MetGlyAlaProThrLeuPro 8
DB 262 ATGGGAGCCCCACACTCCCTCT 285

RESULT 23
US-09-621-976-13420
; Sequence 13420, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 13420
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 255
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-13420

Alignment Scores:
Pred. No.: 7.61 Length: 351
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-621-976-13420 (1-351)
QY 1 MetGlyAlaProThrLeuPro 8
DB 263 ATGGGAGCCCCACACTCCCTCT 286

RESULT 24
US-09-621-976-13407
; Sequence 13407, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 13407
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 254..256
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-13407

Alignment Scores:
Pred. No.: 7.73 Length: 357
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 4 Gaps: 0

DB: 4 Gaps: 0
US-09-690-825-34 (1-142) x US-09-621-976-13407 (1-357)
QY 1 MetGlyAlaProThrLeuPro 8
DB 263 ATGGGAGCCCCACACTCCCTCT 286

RESULT 25
US-09-621-976-13392
; Sequence 13392, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 13392
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-13392

Alignment Scores:
Pred. No.: 7.81 Length: 361
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-621-976-13392 (1-361)
QY 1 MetGlyAlaProThrLeuPro 8
DB 263 ATGGGAGCCCCACACTCCCTCT 286

RESULT 26
US-09-621-976-13428
; Sequence 13428, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 13428
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-13428

Alignment Scores:
Pred. No.: 7.89 Length: 365
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 4 Gaps: 0
```

US-09-690-825-34 (1-142) x US-09-621-976-13428 (1-365)

QY 1 MetGlyAlaProThrLeuProPro 8  
Db 262 ATGGAGGCCCCACACTCCCTCT 285

## RESULT 27

US-09-496-694B-231  
Sequence 231, Application US/09496694B  
Patent No. 6335194  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Elizabeth J. Ackermann  
APPLICANT: Eric E. Swayze  
APPLICANT: Lex M. Cowsett  
TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
FILE REFERENCE: ISPH-0439  
CURRENT APPLICATION NUMBER: US/09/496,694B  
CURRENT FILING DATE: 2000-02-02  
PRIOR APPLICATION NUMBER: 09/286,407  
PRIOR FILING DATE: 1999-04-05  
PRIOR APPLICATION NUMBER: 09/163,162  
PRIOR FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 249  
SEQ ID NO 231  
LENGTH: 515  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 266  
NAME/KEY: CDS  
LOCATION: (301)...(384)  
US-09-496-694B-231

Alignment Scores:  
Pred. No.: 10.9 Length: 515  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: Gaps: 0

US-09-690-825-34 (1-142) x US-09-496-694B-231 (1-515)

QY 122 LysGluPheGluThrAlaLys 129  
Db 325 AAGAGTTTGAAGAGACTGCAGAG 348

## RESULT 28

US-08-119-125A-1  
Sequence 1, Application US/08119125A  
Patent No. 5610011  
GENERAL INFORMATION:  
APPLICANT: SMITH, Hilda Elizabeth  
APPLICANT: VECHT, Uri  
TITLE OF INVENTION: DNA sequences which code for virulence  
TITLE OF INVENTION: Characteristics of Streptococcus suis and parts thereof, polypeptides derived therefrom and the use thereof for the diagnosis and treatment of infection by S. suis in mammals, including protection against infection by S. suis in mammals, including  
TITLE OF INVENTION: antibodies derived therefrom and the use thereof for the diagnosis and treatment of infection by S. suis in mammals, including  
TITLE OF INVENTION: protection against infection by S. suis in mammals, including  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Centraal Diegenesekundig Instituut  
STREET: Fdeltweg 15  
CITY: PH Lelystad  
STATE:  
COUNTRY: The Netherlands  
ZIP: NL-8219  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS v.6.0  
SOFTWARE: WordPerfect v. 6.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,125A  
FILING DATE: 20-SEP-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/NL92/00054  
FILING DATE: 19-MAR-1992  
APPLICATION NUMBER: NL 9100510  
FILING DATE: 21-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Handal, Anthony H.  
REGISTRATION NUMBER: 26275  
REFERENCE/DOCKET NUMBER: SMITHHE119125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 838-8589  
TELEFAX: (203) 838-8794  
INFORMATION FOR SEQ ID NO. 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4376 base pairs  
TYPE: Nucleic acid with corresponding amino acids  
STRANDEDNESS: single stranded  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus suis type II (pathogenic)  
FEATURE:  
OTHER INFORMATION: Extracellular protein factor (EF) gene  
FEATURE:  
NAME/KEY: promoter -35 region  
LOCATION: bp 66 to 71  
FEATURE:  
NAME/KEY: promoter -10 region  
LOCATION: bp 89 to 94  
FEATURE:  
NAME/KEY: promoter -35 region  
LOCATION: bp 153 to 158  
FEATURE:  
NAME/KEY: promoter -10 region  
LOCATION: bp 176 to 181  
FEATURE:  
NAME/KEY: ribosome binding site  
LOCATION: bp 350 to 356  
FEATURE:  
NAME/KEY: signal peptide  
LOCATION: bp 361 to 498  
FEATURE:  
NAME/KEY: mature peptide  
LOCATION: bp 499 to 2890  
FEATURE:  
NAME/KEY: dyad symmetry regions  
LOCATION: from bp 4186 to 4198 and from bp 4203 to 4215  
FEATURE:  
NAME/KEY: dyad symmetry regions  
LOCATION: from bp 4243 to 4257 and from bp 4263 to 4276  
US-08-119-125A-1

Alignment Scores:  
Pred. No.: 78.9 Length: 4376  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.63% Indels: 0  
DB: Gaps: 0

US-09-690-825-34 (1-142) x US-08-119-125A-1 (1-4376)

QY 109 AlalysAsnLysIleAlaLysGlu 116  
Db 2934 GCTAAGATAAGATTGCTAAGAA 2957

## RESULT 29

US-08-119-125A-2  
Sequence 2, Application US/08119125A

```

; Patent No. 5610011
; GENERAL INFORMATION:
; APPLICANT: SMITH, Hilda Elizabeth
; APPLICANT: VECHT, Uri
; TITLE OF INVENTION: DNA Sequences which code for Virulence
; TITLE OF INVENTION: Characteristics of Streptococcus suis and parts thereof, polypeptides derived therefrom and the use thereof for the diagnosis and treatment of infection by S. suis in mammals, including
; TITLE OF INVENTION: antibodies derived therefrom and the use thereof for the diagnosis and treatment of infection by S. suis in mammals, including
; TITLE OF INVENTION: protection against infection by S. suis in mammals, including
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centraal Diergeneeskundig Instituut
; STREET: Edelhertweg 15
; CITY: PH Lelystad
; STATE:
; COUNTRY: The Netherlands
; ZIP: NL-8219
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS v.6.0
; SOFTWARE: Wordperfect v. 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,125A
; FILING DATE: 20-SEP-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL92/00054
; FILING DATE: 19-MAR-1992
; APPLICATION NUMBER: NL 9100510
; FILING DATE: 21-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hangal, Anthony H.
; REGISTRATION NUMBER: 26275
; REFERENCE/DOCKET NUMBER: SMITHHE119125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 838-8589
; TELEFAX: (203) 838-8794
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6744 base pairs
; TYPE: Nucleic acid with corresponding amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus suis type II (pathogenic)
; FEATURE:
; OTHER INFORMATION: Extracellular factor related protein (EF*) gene
; FEATURE:
; NAME/KEY: promoter -35 region
; LOCATION: bp 66 to 71
; FEATURE:
; NAME/KEY: promoter -10 region
; LOCATION: bp 89 to 94
; FEATURE:
; NAME/KEY: promoter -35 region
; LOCATION: bp 153 to 158
; FEATURE:
; NAME/KEY: promoter -10 region
; LOCATION: bp 176 to 181
; FEATURE:
; NAME/KEY: ribosome binding site
; LOCATION: bp 350 to 356
; FEATURE:
; NAME/KEY: signal peptide
; LOCATION: bp 361 to 498
; FEATURE:
; NAME/KEY: start of repetitive units R1-R11
; LOCATION: bp 2869, 3097, 3292, 3520, 4087, 4381, 4609, 4837,
; LOCATION: 5065, 5293, 5521:
; FEATURE:
; NAME/KEY: start of repetitive Asn-Pro-Asn-Leu sequences
; LOCATION: bp 2932, 3160, 3355, 3583, 4150, 4444, 4672, 4900,
; LOCATION: 5128, 5356, 5584:
; FEATURE:
; NAME/KEY: dyad symmetry regions
; LOCATION: from bp 6554 to 6566 and from bp 6571 to 6583
; FEATURE:
; NAME/KEY: dyad symmetry regions
; LOCATION: from bp 6611 to 6625 and from bp 6631 to 6644
; US-08-119-125A-2
; Alignment Scores:
; Pred. No.: 118 Length: 6744
; Score: 8.00 Matches: 8
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 5.63% Indels: 0
; DB: 1 Gaps: 0
; US-09-690-825-34 (1-142) x US-08-119-125A-2 (1-6744)
; QY 109 AlalysAsnLysIleAlaLysGlu 116
; DB 3529 GCTAAGAACAGATTGCCAAGAA 3552
; RESULT 30
; US-09-163-162-2
; Sequence 2, Application US/09163162
; Patent No. 6077709
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Ackermann, Elizabeth J.
; APPLICANT: Swayze, Eric E.
; APPLICANT: Cowsett, Lex M.
; TITLE OF INVENTION: ANTISENSE MODULATION OF Survivin EXPRESSION
; FILE REFERENCE: RTS-0008
; CURRENT APPLICATION NUMBER: US/09/163,162
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 2
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR Primer
; US-09-163-162-2
; Alignment Scores:
; Pred. No.: 6.29 Length: 21
; Score: 7.00 Matches: 7
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 4.93% Indels: 0
; DB: 3 Gaps: 0
; US-09-690-825-34 (1-142) x US-09-163-162-2 (1-21)
; QY 15 LysAspHisArgIleSerThr 21
; DB 1 AAGGACCACCGCATCTCTACA 21
; RESULT 31
; US-09-286-407-2
; Sequence 2, Application US/09286407A
; Patent No. 6165788
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Ackermann, Elizabeth J.
; APPLICANT: Swayze, Eric E.
; APPLICANT: Cowsett, Lex M.
; TITLE OF INVENTION: ANTISENSE MODULATION OF Survivin EXPRESSION
; FILE REFERENCE: ISFH-0349
; CURRENT APPLICATION NUMBER: US/09/286,407A
; CURRENT FILING DATE: 1999-04-05
; NUMBER OF SEQ ID NOS: 48

```

```
; SEQ ID NO 2
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR Primer
US-09-286-407-2

Alignment Scores:
Pred. No.: 6.29 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-286-407-2 (1-21)

QY 15 LysAspHisArgIleSerThr 21
DB 1 AAGGACCACCGCATCTCTACA 21

RESULT 32
US-09-496-694B-4
; Sequence 4, Application US/09496694B
; Patent No. 6335194
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Elizabeth J. Ackermann
; APPLICANT: Eric E. Swayze
; APPLICANT: Lex M. Cowart
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: ISPH-0439
; CURRENT APPLICATION NUMBER: US/09/496,694B
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 09/286,407
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 09/163,162
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 249
; SEQ ID NO 4
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR Primer
US-09-496-694B-4

Alignment Scores:
Pred. No.: 6.29 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-496-694B-4 (1-21)

QY 15 LysAspHisArgIleSerThr 21
DB 1 AAGGACCACCGCATCTCTACA 21

RESULT 33
US-09-163-162-3/c
; Sequence 3, Application US/09163162
; Patent No. 607709
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Ackermann, Elizabeth J.
; APPLICANT: Swayze, Eric E.
; APPLICANT: Cowart, Lex M.
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: RTS-0008
; CURRENT APPLICATION NUMBER: US/09/163,162
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 3
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR Primer
US-09-163-162-3

Alignment Scores:
Pred. No.: 6.57 Length: 22
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-163-162-3 (1-22)

QY 48 ThrGluAsnGluProAspLeu 54
DB 22 ACTGAGAACGAGCCAGACTTG 2

RESULT 34
US-09-286-407-3/c
; Sequence 3, Application US/09286407A
; Patent No. 6165788
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Ackermann, Elizabeth J.
; APPLICANT: Swayze, Eric E.
; APPLICANT: Cowart, Lex M.
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: ISPH-0349
; CURRENT APPLICATION NUMBER: US/09/286,407A
; CURRENT FILING DATE: 1999-04-05
; NUMBER OF SEQ ID NOS: 48
; SEQ ID NO 3
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR Primer
US-09-286-407-3

Alignment Scores:
Pred. No.: 6.57 Length: 22
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-286-407-3 (1-22)

QY 48 ThrGluAsnGluProAspLeu 54
DB 22 ACTGAGAACGAGCCAGACTTG 2

RESULT 35
US-09-496-694B-5/c
; Sequence 5, Application US/09496694B
; Patent No. 6335194
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Elizabeth J. Ackermann
; APPLICANT: Eric E. Swayze
; APPLICANT: Lex M. Cowart
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: ISPH-0439
; CURRENT APPLICATION NUMBER: US/09/496,694B
```

; CURRENT FILING DATE: 2000-02-02  
 ; PRIOR APPLICATION NUMBER: 09/286,407  
 ; PRIOR FILING DATE: 1999-04-05  
 ; PRIOR APPLICATION NUMBER: 09/163,162  
 ; PRIOR FILING DATE: 1998-09-29  
 ; NUMBER OF SEQ ID NOS: 249  
 ; SEQ ID NO 5  
 ; LENGTH: 22  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: PCR Primer  
 US-09-496-694B-5

Alignment Scores:  
 Pred. No.: 6.57 Length: 22  
 Score: 7.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.93% Indels: 0  
 DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-496-694B-5 (1-22)

QY 48 ThrGluAsnGluProAspLeu 54  
 DB 22 ACTGAGACGAGCCAGACTTG 2

## RESULT 36

US-09-163-162-4  
 ; Sequence 4, Application US/09163162  
 ; Patent No. 6077709  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bennett, C. Frank  
 ; APPLICANT: Ackermann, Elizabeth J.  
 ; APPLICANT: Swayze, Eric E.  
 ; APPLICANT: Cowsett, Lex M.  
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
 ; FILE REFERENCE: RTS-0008  
 ; CURRENT APPLICATION NUMBER: US/09/163,162  
 ; CURRENT FILING DATE: 1998-09-29  
 ; NUMBER OF SEQ ID NOS: 47  
 ; SEQ ID NO 4  
 ; LENGTH: 23  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: PCR Probe  
 US-09-163-162-4

Alignment Scores:  
 Pred. No.: 6.84 Length: 23  
 Score: 7.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.93% Indels: 0  
 DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-163-162-4 (1-23)

QY 40 GluAlaGlyPheIleHisCys 46  
 DB 2 GAGCGTGGCTTCATCCACTGC 22

## RESULT 37

US-09-286-407-4  
 ; Sequence 4, Application US/09286407A  
 ; Patent No. 6165788  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bennett, C. Frank  
 ; APPLICANT: Ackermann, Elizabeth J.  
 ; APPLICANT: Swayze, Eric E.  
 ; APPLICANT: Cowsett, Lex M.

; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
 ; FILE REFERENCE: ISPH-0349  
 ; CURRENT APPLICATION NUMBER: US/09/286,407A  
 ; CURRENT FILING DATE: 1999-04-05  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SEQ ID NO 4  
 ; LENGTH: 23  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: PCR Probe  
 US-09-286-407-4

Alignment Scores:  
 Pred. No.: 6.84 Length: 23  
 Score: 7.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.93% Indels: 0  
 DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-286-407-4 (1-23)

QY 40 GluAlaGlyPheIleHisCys 46  
 DB 2 GAGCGTGGCTTCATCCACTGC 22

## RESULT 38

US-09-496-694B-6  
 ; Sequence 6, Application US/09496694B  
 ; Patent No. 6335194  
 ; GENERAL INFORMATION:  
 ; APPLICANT: C. Frank Bennett  
 ; APPLICANT: Elizabeth J. Ackermann  
 ; APPLICANT: Eric E. Swayze  
 ; APPLICANT: Lex M. Cowsett  
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
 ; FILE REFERENCE: ISPH-0439  
 ; CURRENT APPLICATION NUMBER: US/09/496,694B  
 ; CURRENT FILING DATE: 2000-02-02  
 ; PRIOR APPLICATION NUMBER: 09/286,407  
 ; PRIOR FILING DATE: 1999-04-05  
 ; PRIOR APPLICATION NUMBER: 09/163,162  
 ; PRIOR FILING DATE: 1998-09-29  
 ; NUMBER OF SEQ ID NOS: 249  
 ; SEQ ID NO 6  
 ; LENGTH: 23  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: PCR Probe  
 US-09-496-694B-6

Alignment Scores:  
 Pred. No.: 6.84 Length: 23  
 Score: 7.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.93% Indels: 0  
 DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-496-694B-6 (1-23)

QY 40 GluAlaGlyPheIleHisCys 46  
 DB 2 GAGCGTGGCTTCATCCACTGC 22

## RESULT 39

US-09-313-294A-1818  
 ; Sequence 1818, Application US/09313294A  
 ; Patent No. 6476212  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lalgudi, Raghunath V.



APPLICANT: Ito, Laura Y.  
APPLICANT: Sherman, Bradley K.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
FILE REFERENCE: PL-0017 US  
CURRENT APPLICATION NUMBER: US/09/313,294A  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 7600  
SOFTWARE: PERL Program  
SEQ ID NO 1818  
LENGTH: 273  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6476212 700551571H1  
US-09-313-294A-1818

Alignment Scores:  
Pred. No.: 67.8 Length: 273  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.93% Indels: 0  
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-313-294A-1818 (1-273)

Qy 104 Leuaspargluargalalys 110  
Db 9 CTAGATCGGAGCGCGCAAG 29

## RESULT 40

US-09-328-111-715/c  
Sequence 715, Application US/09328111  
Patent No. 6262333  
GENERAL INFORMATION:  
APPLICANT: Endege, Wilson O.  
APPLICANT: Steinmann, Kathleen E.  
APPLICANT: Astle, Jon H.  
APPLICANT: Burgess, Christopher C.  
APPLICANT: Bushnell, Steven E.  
APPLICANT: Carroll III, Eddie  
APPLICANT: Catino, Theodore J.  
APPLICANT: Derti, Adnan  
APPLICANT: Ford, Donna M.  
APPLICANT: Lewis, Marcia E.  
APPLICANT: Monahan, John E.  
APPLICANT: Schlegel, Robert  
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
FILE REFERENCE: CCD-257 (US)  
CURRENT APPLICATION NUMBER: US/09/328,111  
CURRENT FILING DATE: 1999-06-08  
EARLIER APPLICATION NUMBER: US 60/088,801  
EARLIER FILING DATE: 1998-06-10  
NUMBER OF SEQ ID NOS: 850  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 715  
LENGTH: 310  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1) - (310)  
OTHER INFORMATION: n = A, T, C or G  
US-09-328-111-715

Alignment Scores:  
Pred. No.: 76.2 Length: 310  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.93% Indels: 0

DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-328-111-715 (1-310)

Qy 14 LeuLysAspHisArgIleSer 20  
Db 74 TTAAGATCACAGATCTCT 54

## RESULT 41

US-09-107-532A-1678  
Sequence 1678, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 1678:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 504 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...504  
SEQUENCE DESCRIPTION: SEQ ID NO: 1678:

US-09-107-532A-1678

Alignment Scores:

Pred. No.: 120 Length: 504  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.93% Indels: 0  
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-107-532A-1678 (1-504)

Qy 131 ValArgGluAlaIleGluGln 137

Db 433 GTAAGACGTGCAATTGAGCAA 453

RESULT 42

US-09-621-976-10449

; Sequence 10449, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Robert S.

; APPLICANT: Giordano J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054P2

; CURRENT APPLICATION NUMBER: US/09/621.976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 10449

; LENGTH: 559

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-621-976-10449

Alignment Scores:

Pred. No.:	132	Length:	559
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.93%	Indels:	0
DB:	4	Gaps:	0

US-09-690-825-34 (1-142) x US-09-621-976-10449 (1-559)

Qy 104 LeuAspArgGluArgAlaLys 110

Db 37 CTGGACAGGGAGAGGGCCAAA 57

RESULT 43

US-09-107-532A-1373

; Sequence 1373, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; NUMBER OF SEQUENCES: 710

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 1373:

; SEQUENCE CHARACTERISTICS:

LENGTH: 561 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...561

SEQUENCE DESCRIPTION: SEQ ID NO: 1373:

US-09-107-532A-1373

Alignment Scores:

Pred. No.:	132	Length:	561
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.93%	Indels:	0
DB:	4	Gaps:	0

US-09-690-825-34 (1-142) x US-09-107-532A-1373 (1-561)

Qy 127 ThrAlaLysValArgArg 133

Db 127 ACTGCAGAAAGTACGACGT 147

RESULT 44

US-09-252-991A-4396

; Sequence 4396, Application US/09252991A

; Patent No. 6551755

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 4396

; LENGTH: 588

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-4396

Alignment Scores:

Pred. No.:	138	Length:	588
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.93%	Indels:	0
DB:	4	Gaps:	0

US-09-690-825-34 (1-142) x US-09-252-991A-4396 (1-588)

Qy 6 LeuProProAlaTrpGlnPro 12

Db 30 TTGCCACCGCGCTGGCAGCCT 50

RESULT 45

US-08-642-274D-15/c

; Sequence 16, Application US/08642274D

; Patent No. 6200749

; GENERAL INFORMATION:

; APPLICANT: Shiloh, Yosef

; TITLE OF INVENTION: MUTATED FORMS OF THE ATAXIA-TELANGIECTASIA GENE AND METHOD TO

; TITLE OF INVENTION: SCREEN FOR A PARTIAL A-T PHENOTYPE
; FILE REFERENCE: 229000033
; CURRENT APPLICATION NUMBER: US/08/642,274D
; CURRENT FILING DATE: 1996-05-03
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: genomic
US-08-642-274D-16

Alignment Scores:
Pred. No.: 143 Length: 611
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-08-642-274D-16 (1-611)
Qy 86 PheLeuSerVallysLysGln 92
Db 222 TTCTTGTCTGTGAAGAACAA 202

RESULT 46
US-08-952-014C-16/c
; Sequence 16, Application US/08952014C
; Patent No. 6265158
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE AND ITS
; TITLE OF INVENTION: GENOMIC ORGANIZATION
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 6265158thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/952,014C
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,995
; REFERENCE/DOCKET NUMBER: 2290.00028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 810-539-5050
; TELEFAX: 810-539-5055
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 611 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-952-014C-16

Alignment Scores:
Pred. No.: 143 Length: 611
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-08-952-014C-16 (1-611)
Qy 86 PheLeuSerVallysLysGln 92
Db 222 TTCTTGTCTGTGAAGAACAA 202

RESULT 47
US-09-134-000C-912
; Sequence 912, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 912
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-912

Alignment Scores:
Pred. No.: 146 Length: 627
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-134-000C-912 (1-627)
Qy 127 ThrAlaLysLysValArgArg 133
Db 193 ACGGCCAAAAAAGTGGCGCGT 213

RESULT 48
US-09-489-039A-6367/c
; Sequence 6367, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 6367
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6367

Alignment Scores:
Pred. No.: 150 Length: 642
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-489-039A-6367 (1-642)

QY 3 AlaProThrLeuProProAla 9  
 Db 274 GCGCCACCCCTCCGCGCTGCA 254

# RESULT 49

US-09-134-000C-375  
 ; Sequence 375, Application US/09134000C  
 ; Patent No. 6617156  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 032796-032  
 ; CURRENT APPLICATION NUMBER: US/09/134,000C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/055,778  
 ; PRIOR FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 6812  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 375  
 ; LENGTH: 777  
 ; TYPE: DNA  
 ; ORGANISM: Enterococcus faecalis  
 US-09-134-000C-375

Alignment Scores:  
 Pred. No.: 179 Length: 777  
 Score: 7.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.93% Indels: 0  
 DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-134-000C-375 (1-777)

QY 125 GluThrAlaLysVal 131  
 Db 511 GAAGAACGGCTAAAAAAGTG 531

# RESULT 50

US-09-252-991A-4263/c  
 ; Sequence 4263, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 4263  
 ; LENGTH: 924  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-4263

Alignment Scores:  
 Pred. No.: 210 Length: 924  
 Score: 7.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.93% Indels: 0  
 DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-252-991A-4263 (1-924)

QY 6 LeuProProAlaTrpGlnPro 12  
 Db 915 TTGCCACCGCGCTGGCGCCT 895

# RESULT 51

US-09-252-991A-3124/c  
 ; Sequence 3124, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 3124  
 ; LENGTH: 939  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-3124

Alignment Scores:  
 Pred. No.: 213 Length: 939  
 Score: 7.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.93% Indels: 0  
 DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-252-991A-3124 (1-939)

QY 30 GlyCysAlaCysThrProGlu 36  
 Db 915 GGTTCGCATGCACGCTGAG 895

# RESULT 52

US-09-475-316A-117  
 ; Sequence 117, Application US/09475316A  
 ; Patent No. 6210942  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lewis, No. 6210942man G.  
 ; APPLICANT: Davin, Laurence B.  
 ; APPLICANT: Dinkova-Kostova, Albena T.  
 ; APPLICANT: Fujita, Masayuki  
 ; APPLICANT: Gang, David R.  
 ; APPLICANT: Sarxanen, Simo  
 ; APPLICANT: Ford, Joshua D  
 ; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASES,  
 ; FILE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE  
 ; FILE REFERENCE: WSUR-1-13793  
 ; CURRENT APPLICATION NUMBER: US/09/475,316A  
 ; CURRENT FILING DATE: 1999-12-30  
 ; PRIOR APPLICATION NUMBER: 09/307,653  
 ; PRIOR FILING DATE: 1999-05-07  
 ; PRIOR APPLICATION NUMBER: PCT/US97/20391  
 ; PRIOR FILING DATE: 1997-11-07  
 ; PRIOR APPLICATION NUMBER: 60/054,380  
 ; PRIOR FILING DATE: 1997-07-31  
 ; PRIOR APPLICATION NUMBER: 60/030,522  
 ; PRIOR FILING DATE: 1996-11-08  
 ; NUMBER OF SEQ ID NOS: 122  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 117  
 ; LENGTH: 942  
 ; TYPE: DNA  
 ; ORGANISM: Schisandra chinensis  
 ; FEATURE:  
 ; NAME/KEY: CDS

```

; LOCATION: (1)..(942)
US-09-475-316A-117

Alignment Scores:
Pred. No.: 214 Length: 942
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-475-316A-117 (1-942)

QY 130 LysValArgAlaIleGlu 136
Db 424 AAGGTCAGCGTGCATAGAA 444

RESULT 53
US-09-704-640-117
; Sequence 117, Application US/09704640
; Patent No. 6635459
; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6635459man G.
; APPLICANT: Davin, Laurence B.
; APPLICANT: Dinkova-Kostova, Albena T.
; APPLICANT: Fujita, Masayuki
; APPLICANT: Gang, David R.
; APPLICANT: Sarkanen, Simo
; APPLICANT: Ford, Joshua D
; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASE,
; FILE REFERENCE: WSUR-1-16492
; CURRENT FILING DATE: 1999-12-30
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 09/307,653
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/054,380
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: 60/030,522
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 117
; TYPE: DNA
; LENGTH: 942
; ORGANISM: Schisandra chinensis
US-09-475-316A-119

Alignment Scores:
Pred. No.: 214 Length: 944
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-475-316A-119 (1-944)

QY 130 LysValArgAlaIleGlu 136
Db 424 AAGGTCAGCGTGCATAGAA 444

RESULT 55
US-09-704-640-119
; Sequence 119, Application US/09704640
; Patent No. 6635459
; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6635459man G.
; APPLICANT: Davin, Laurence B.
; APPLICANT: Dinkova-Kostova, Albena T.
; APPLICANT: Fujita, Masayuki
; APPLICANT: Gang, David R.
; APPLICANT: Sarkanen, Simo
; APPLICANT: Ford, Joshua D
; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASE,
; FILE REFERENCE: WSUR-1-16492
; CURRENT FILING DATE: 1999-12-30
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 09/307,653
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/054,380
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: 60/030,522
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 119
; TYPE: DNA
; LENGTH: 944
; ORGANISM: Schisandra chinensis
US-09-475-316A-119

Alignment Scores:
Pred. No.: 214 Length: 942
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-704-640-117 (1-942)

QY 130 LysValArgAlaIleGlu 136
Db 424 AAGGTCAGCGTGCATAGAA 444

RESULT 54
US-09-475-316A-119
; Sequence 119, Application US/09475316A
; Patent No. 6210942
```

```

; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6210942man G.
; APPLICANT: Davin, Laurence B.
; APPLICANT: Dinkova-Kostova, Albena T.
; APPLICANT: Fujita, Masayuki
; APPLICANT: Gang, David R.
; APPLICANT: Sarkanen, Simo
; APPLICANT: Ford, Joshua D
; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASES,
; FILE REFERENCE: WSUR-1-13793
; CURRENT FILING DATE: 1999-12-30
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 09/307,653
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/054,380
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: 60/030,522
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 119
; TYPE: DNA
; LENGTH: 944
; ORGANISM: Schisandra chinensis
US-09-475-316A-119

Alignment Scores:
Pred. No.: 214 Length: 944
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-475-316A-119 (1-944)

QY 130 LysValArgAlaIleGlu 136
Db 424 AAGGTCAGCGTGCATAGAA 444

RESULT 55
US-09-704-640-119
; Sequence 119, Application US/09704640
; Patent No. 6635459
; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6635459man G.
; APPLICANT: Davin, Laurence B.
; APPLICANT: Dinkova-Kostova, Albena T.
; APPLICANT: Fujita, Masayuki
; APPLICANT: Gang, David R.
; APPLICANT: Sarkanen, Simo
; APPLICANT: Ford, Joshua D
; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASE,
; FILE REFERENCE: WSUR-1-16492
; CURRENT FILING DATE: 2000-11-02
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/475,316
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/054,380
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/054,380
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: 60/030,522
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 119
```

```

; LENGTH: 944
; TYPE: DNA
; ORGANISM: Schisandra chinensis
US-09-704-640-119

Alignment Scores:
Pred. No.: 214 Length: 944
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-704-640-119 (1-944)

QY 130 LysValArgAlaIleGlu 136
Db 424 AAGGTCAGCGTGAATAGAA 444

RESULT 56
US-09-489-039A-5085
; Sequence 5085 Application US/09489039A
; Patent No. 6810836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5085
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5085

Alignment Scores:
Pred. No.: 215 Length: 951
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-489-039A-5085 (1-951)

QY 5 ThrLeuProAlaTyrGln 11
Db 587 ACGCTACCGCAGCATGCCAA 607

RESULT 57
US-08-969-644-21/c
; Sequence 21, Application US/08969644
; Patent No. 6096519
; GENERAL INFORMATION:
; APPLICANT: Ratti, Giulio
; APPLICANT: Comanducci, Maurizio
; APPLICANT: Tecce, Mario F.
; APPLICANT: Giuliani, Marzia M.
; TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA
; TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY
; TITLE OF INVENTION: THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID
; TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 N. Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA

; ZIP: 22046-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/969,644
; APPLICATION NUMBER: US/08/969,644
; FILING DATE: 13-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/467,152
; FILING DATE:
; APPLICATION NUMBER: US/07/661,820
; FILING DATE:
; APPLICATION NUMBER: IT MI 91A000314
; FILING DATE: 07-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 1267-202P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Chlamydia trachomatis
; STRAIN: GO/86 serotype D (trachoma biovar)
; IMMEDIATE SOURCE:
; CLONE: pUC8-pCO plasmid, ATCC 6814
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..993
US-08-969-644-21

Alignment Scores:
Pred. No.: 224 Length: 993
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-08-969-644-21 (1-993)

QY 98 LeuGlyGlupheLeuLysLeu 104
Db 88 CTGGCGGATTTTAAACTT 68

RESULT 58
US-08-444-189-21/c
; Sequence 21, Application US/08444189
; Patent No. 6110705
; GENERAL INFORMATION:
; APPLICANT: Ratti, Giulio
; APPLICANT: Comanducci, Maurizio
; APPLICANT: Tecce, Mario F.
; APPLICANT: Giuliani, Marzia M.
; TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA
; TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY
; TITLE OF INVENTION: THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID
; TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 N. Washington Street

```

CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,189  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/180,528  
FILING DATE:  
APPLICATION NUMBER: US/07/991,512  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820  
FILING DATE:  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORGANISM: Chlamydia trachomatis  
STRAIN: GO/86 serotype D (trachoma biovar)  
IMMEDIATE SOURCE:  
CLONE: pUC8-pGO plasmid, ATCC 68314  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..993  
US-08-444-189-21

Alignment Scores:  
Pred. No.: 224 Length: 993  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.93% Indels: 0  
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-08-444-189-21 (1-993)  
QY 98 LeuGlyGluPheLeuLysLeu 104  
Db 88 CTGGCGAATTTTAAACTT 68

RESULT 59  
US-08-468-544-21/c  
Sequence 21, Application US/08468544  
Patent No. 6248563  
GENERAL INFORMATION:  
APPLICANT: Ratti, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID

TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,544  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/661,820  
FILING DATE: 28-FEB-1991  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORGANISM: Chlamydia trachomatis  
STRAIN: GO/86 serotype D (trachoma biovar)  
IMMEDIATE SOURCE:  
CLONE: pUC8-pGO plasmid, ATCC 68314  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..993  
US-08-468-544-21

Alignment Scores:  
Pred. No.: 224 Length: 993  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.93% Indels: 0  
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-08-468-544-21 (1-993)  
QY 98 LeuGlyGluPheLeuLysLeu 104  
Db 88 CTGGCGAATTTTAAACTT 68

RESULT 60  
US-08-978-404B-9  
Sequence 9, Application US/08978404B  
Patent No. 5968782  
GENERAL INFORMATION:  
APPLICANT: Stevens, Richard L.  
TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
TITLE OF INVENTION: FIBRINOGEN  
NUMBER OF SEQUENCES: 74  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

```
/ STREET: 600 Atlantic Avenue
/ CITY: Boston
/ STATE: MA
/ COUNTRY: U.S.A.
/ ZIP: 02210-2211
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: Fast-Seq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/978,404B
/ FILING DATE: 25-NOV-97
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/032,354
/ FILING DATE: 04-DEC-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Plumer, Elizabeth R.
/ REGISTRATION NUMBER: 36,637
/ REFERENCE/DOCKET NUMBER: B0801/7090
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-720-3500
/ TELEFAX: 617-720-2441
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1095 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-978-404B-9

Alignment Scores:
Pred. No.: 246 Length: 1095
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 2 Gaps: 0

US-09-690-825-34 (1-142) x US-08-978-404B-9 (1-1095)

QY 3 AlaProThrLeuProAla 9
Db 646 GCGCCACCTTACCACCTGCA 666

RESULT 61
US-09-475-316A-61
; Sequence 61, Application US/09475316A
; Patent No. 6210942
; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6210942man G.
; APPLICANT: Davin, Laurence B.
; APPLICANT: Dinkova-Kostova, Albena T.
; APPLICANT: Fujita, Masayuki
; APPLICANT: Gang, David R.
; APPLICANT: Sarkanen, Simo
; APPLICANT: Ford, Joshua D
; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASES,
; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE
; FILE REFERENCE: WSUR-1-13793
; CURRENT APPLICATION NUMBER: US/09/475,316A
; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/307,653
; PRIOR FILING DATE: 1999-05-07
; PRIOR FILING DATE: 1997-11-07
; PRIOR FILING DATE: 1997-07-31
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 1190
; TYPE: DNA
; ORGANISM: Thuja plicata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(951)
; US-09-704-640-61

Alignment Scores:
Pred. No.: 265 Length: 1190
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 4 Gaps: 0
```

```
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 61
/ LENGTH: 1190
/ TYPE: DNA
/ ORGANISM: Thuja plicata
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (13)..(951)
/ US-09-475-316A-61

Alignment Scores:
Pred. No.: 265 Length: 1190
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-475-316A-61 (1-1190)

QY 130 LysValArgAlaIleGlu 136
Db 430 AAGGTCGGCGTGCCATTGAA 450

RESULT 62
US-09-704-640-61
; Sequence 61, Application US/09704640
; Patent No. 6635459
; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6635459man G.
; APPLICANT: Davin, Laurence B.
; APPLICANT: Dinkova-Kostova, Albena T.
; APPLICANT: Fujita, Masayuki
; APPLICANT: Gang, David R.
; APPLICANT: Sarkanen, Simo
; APPLICANT: Ford, Joshua D
; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASE,
; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEIN AND METHODS OF USE
; FILE REFERENCE: WSUR-1-16492
; CURRENT APPLICATION NUMBER: US/09/704,640
; CURRENT FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 09/475,316
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/307,653
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: PCT/US97/20391
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/054,380
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: 60/030,522
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 1190
; TYPE: DNA
; ORGANISM: Thuja plicata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(951)
; US-09-704-640-61

Alignment Scores:
Pred. No.: 265 Length: 1190
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.93% Indels: 0
DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-704-640-61 (1-1190)

QY 130 LysValArgAlaIleGlu 136
```



```
Db      430  AAGTTCGGCGTGCATTGAA 450
|||||
RESULT 63
US-09-328-352-3331/c
; Sequence 3331, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Berton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GT999-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 331
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3331
Alignment Scores:
Pred. No.: 266      Length: 1194
Score: 7.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 4.93%      Indels: 0
DB: 4      Gaps: 0

US-09-690-825-34 (1-142) x US-09-328-352-3331 (1-1194)
QY      109  AlaLysAsnLysIleAlaLys 115
|||||
Db      473  GCTAAAAACAAATAGCAAAA 453
|||||

RESULT 64
US-09-252-991A-8373/c
; Sequence 8373, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8373
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8373
Alignment Scores:
Pred. No.: 273      Length: 1230
Score: 7.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 4.93%      Indels: 0
DB: 4      Gaps: 0

US-09-690-825-34 (1-142) x US-09-252-991A-8373 (1-1230)
QY      8  ProAlatrpGlnPropheLeu 14
|||||
Db      99  CCGGCGTGGCAGCGTTCCTC 79
|||||

RESULT 65
US-09-252-991A-8504
; Sequence 8504, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8504
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8504
Alignment Scores:
Pred. No.: 279      Length: 1257
Score: 7.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 4.93%      Indels: 0
DB: 4      Gaps: 0

US-09-690-825-34 (1-142) x US-09-252-991A-8504 (1-1257)
QY      8  ProAlatrpGlnPropheLeu 14
|||||
Db      1216  CCGGCGTGGCAGCGTTCCTC 1236
|||||

RESULT 66
US-09-475-316A-69
; Sequence 69, Application US/09475316A
; Patent No. 6210942
; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6210942man G.
; APPLICANT: Davin, Laurence B.
; APPLICANT: Dinkova-kostova, Albena T.
; APPLICANT: Fujita, Masayuki
; APPLICANT: Gang, David R.
; APPLICANT: Sarkanen, Simo
; APPLICANT: Ford, Joshua D
; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASES,
; FILE REFERENCE: WSUR-1-13793
; CURRENT APPLICATION NUMBER: US/09/475,316A
; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/307,653
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: PCT/US97/20391
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/054,380
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: 60/030,522
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 69
; LENGTH: 1282
; TYPE: DNA
; ORGANISM: Tsuga heterophylla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(922)
US-09-475-316A-69
Alignment Scores:
Pred. No.: 284      Length: 1282
Score: 7.00      Matches: 7
```

Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.93% Indels: 0  
 DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-475-316A-69 (1-1282)

QY 130 LysValArgAlaIleGlu 136  
 DB 404 AAGGTCGGCGCCATTGAA 424

RESULT 67

US-09-704-640-69  
 ; Sequence 69, Application US/09704640  
 ; Patent No. 6635459

; GENERAL INFORMATION:

; APPLICANT: Lewis, No. 6635459man G.  
 ; APPLICANT: Davin, Laurence B.  
 ; APPLICANT: Dinkova-Kostova, Albena T.  
 ; APPLICANT: Fujita, Masayuki  
 ; APPLICANT: Gang, David R.  
 ; APPLICANT: Sarkanen, Simo  
 ; APPLICANT: Ford, Joshua D.

; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASE,  
 ; FILE REFERENCE: WSUR-1-16492

; CURRENT APPLICATION NUMBER: US/09/704,640  
 ; PRIOR FILING DATE: 2000-11-02

; PRIOR APPLICATION NUMBER: 09/475,316  
 ; PRIOR FILING DATE: 1999-12-30

; PRIOR APPLICATION NUMBER: 09/307,653  
 ; PRIOR FILING DATE: 1999-05-07

; PRIOR APPLICATION NUMBER: PCT/US97/20391  
 ; PRIOR FILING DATE: 1997-11-07

; PRIOR APPLICATION NUMBER: 60/054,380  
 ; PRIOR FILING DATE: 1997-07-31

; PRIOR APPLICATION NUMBER: 60/030,522  
 ; PRIOR FILING DATE: 1996-11-08

; NUMBER OF SEQ ID NOS: 122  
 ; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 69  
 ; LENGTH: 1282

; TYPE: DNA  
 ; ORGANISM: Tsuga heterophylla

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (2)..(922)

US-09-704-640-69

Alignment Scores:  
 Pred. No.: 284 Length: 1282  
 Score: 7.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.93% Indels: 0  
 DB: 4 Gaps: 0

US-09-690-825-34 (1-142) x US-09-704-640-69 (1-1282)

QY 130 LysValArgAlaIleGlu 136  
 DB 404 AAGGTCGGCGCCATTGAA 424

RESULT 68

US-09-475-316A-65

; Sequence 65, Application US/09475316A  
 ; Patent No. 6210942

; GENERAL INFORMATION:

; APPLICANT: Lewis, No. 6210942man G.  
 ; APPLICANT: Davin, Laurence B.  
 ; APPLICANT: Dinkova-Kostova, Albena T.  
 ; APPLICANT: Fujita, Masayuki  
 ; APPLICANT: Gang, David R.

; APPLICANT: Sarkanen, Simo  
 ; APPLICANT: Ford, Joshua D  
 ; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASE,  
 ; FILE REFERENCE: WSUR-1-13793

; CURRENT APPLICATION NUMBER: US/09/475,316A  
 ; CURRENT FILING DATE: 1999-12-30

; PRIOR APPLICATION NUMBER: 09/307,653  
 ; PRIOR FILING DATE: 1999-05-07

; PRIOR APPLICATION NUMBER: PCT/US97/20391  
 ; PRIOR FILING DATE: 1997-11-07

; PRIOR APPLICATION NUMBER: 60/054,380  
 ; PRIOR FILING DATE: 1997-07-31

; PRIOR APPLICATION NUMBER: 60/030,522  
 ; PRIOR FILING DATE: 1996-11-08

; NUMBER OF SEQ ID NOS: 122  
 ; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 65

; LENGTH: 1308

; TYPE: DNA

; ORGANISM: Thuja plicata

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (164)..(1105)

US-09-475-316A-65

Alignment Scores:

Pred. No.: 289 Length: 1308  
 Score: 7.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.93% Indels: 0  
 DB: 3 Gaps: 0

US-09-690-825-34 (1-142) x US-09-475-316A-65 (1-1308)

QY 130 LysValArgAlaIleGlu 136

DB 581 AAGGTCGGCGCCATTGAA 601

RESULT 69

US-09-704-640-65

; Sequence 65, Application US/09704640

; Patent No. 6635459

; GENERAL INFORMATION:

; APPLICANT: Lewis, No. 6635459man G.

; APPLICANT: Davin, Laurence B.

; APPLICANT: Dinkova-Kostova, Albena T.

; APPLICANT: Fujita, Masayuki

; APPLICANT: Gang, David R.

; APPLICANT: Sarkanen, Simo

; APPLICANT: Ford, Joshua D

; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASE,  
 ; FILE REFERENCE: WSUR-1-16492

; CURRENT APPLICATION NUMBER: US/09/704,640  
 ; CURRENT FILING DATE: 2000-11-02

; PRIOR APPLICATION NUMBER: 09/475,316  
 ; PRIOR FILING DATE: 1999-12-30

; PRIOR APPLICATION NUMBER: 09/307,653  
 ; PRIOR FILING DATE: 1999-05-07

; PRIOR APPLICATION NUMBER: PCT/US97/20391  
 ; PRIOR FILING DATE: 1997-11-07

; PRIOR APPLICATION NUMBER: 60/054,380  
 ; PRIOR FILING DATE: 1997-07-31

; PRIOR APPLICATION NUMBER: 60/030,522  
 ; PRIOR FILING DATE: 1996-11-08

; NUMBER OF SEQ ID NOS: 122  
 ; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 65

; LENGTH: 1308

; TYPE: DNA

; ORGANISM: Thuja plicata

FEATURE:	NAME/KEY:	CDS	LOCATION:	(164)...	(1105)
US-09-704-640-65					
Alignment Scores:		299	Length:	1308	
Pred. No.:		7	Matches:	7	
Score:		100.00%	Conservative:	0	
Percent Similarity:		100.00%	Mismatches:	0	
Best Local Similarity:		100.00%	Indels:	0	
Query Match:		4.93%	Gaps:	0	
DB:		4			
US-09-690-825-34 (1-142) x US-09-704-640-65 (1-1308)					
QY	130	LysValArgAlaIleGlu 136			
DB	581	AAAGTTAGACGTGCCATTGAA 601			
RESULT 70					
US-09-475-316A-71					
; Sequence 71, Application US/09475316A					
; Patent No. 6210942					
GENERAL INFORMATION:					
; APPLICANT: Lewis, No. 6210942man G.					
; APPLICANT: Davin, Laurence B.					
; APPLICANT: Dinkova-Kostova, Albena T.					
; APPLICANT: Fujita, Masayuki					
; APPLICANT: Gang, David R.					
; APPLICANT: Sarkanen, Simo					
; APPLICANT: Ford, Joshua D					
TITLE OF INVENTION: RECOMBINANT PINOESINOL/LARICINESINOL REDUCTASES,					
FILE REFERENCE: WSUR-1-16492					
CURRENT APPLICATION NUMBER: US/09/704,640					
CURRENT FILING DATE: 2000-11-02					
PRIOR APPLICATION NUMBER: 09/475,316					
PRIOR FILING DATE: 1999-12-30					
PRIOR APPLICATION NUMBER: 09/307,653					
PRIOR FILING DATE: 1999-05-07					
PRIOR APPLICATION NUMBER: PCT/US97/20391					
PRIOR FILING DATE: 1997-11-07					
PRIOR APPLICATION NUMBER: 60/054,380					
PRIOR FILING DATE: 1997-07-31					
PRIOR APPLICATION NUMBER: 60/030,522					
PRIOR FILING DATE: 1996-11-08					
NUMBER OF SEQ ID NOS: 122					
SOFTWARE: PatentIn Ver. 2.0					
SEQ ID NO 71					
; LENGTH: 1328					
; TYPE: DNA					
; ORGANISM: Tsuga heterophylla					
; FEATURE:					
; NAME/KEY: CDS					
; LOCATION: (20)..(946)					
US-09-704-640-71					
Alignment Scores:					
Pred. No.:	294	Length:	1328		
Score:	7.00	Matches:	7		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	4.93%	Indels:	0		
DB:	4	Gaps:	0		
US-09-690-825-34 (1-142) x US-09-704-640-71 (1-1328)					
QY	130	LysValArgAlaIleGlu 136			
DB	428	AAGGTCGGCGCCATTGAA 448			
RESULT 72					
US-09-205-258-212					
; Sequence 212, Application US/09205258					
; Patent No. 6525174					
GENERAL INFORMATION:					
; APPLICANT: Young et al.					
TITLE OF INVENTION: 207 Human Secreted Proteins					
FILE REFERENCE: PZ00721					
CURRENT APPLICATION NUMBER: US/09/205,258					
CURRENT FILING DATE: 1998-12-04					
EARLIER APPLICATION NUMBER: PCT/US98/11422					
EARLIER FILING DATE: 1998-06-04					
EARLIER APPLICATION NUMBER: 60/048,885					
EARLIER FILING DATE: 1997-06-06					
EARLIER APPLICATION NUMBER: 60/049,375					
EARLIER FILING DATE: 1997-06-06					
EARLIER APPLICATION NUMBER: 60/048,881					
EARLIER FILING DATE: 1997-06-06					
EARLIER APPLICATION NUMBER: 60/048,880					
EARLIER FILING DATE: 1997-06-06					
EARLIER APPLICATION NUMBER: 60/048,896					
EARLIER FILING DATE: 1997-06-06					
EARLIER APPLICATION NUMBER: 60/049,020					
EARLIER FILING DATE: 1997-06-06					
EARLIER APPLICATION NUMBER: 60/048,876					
US-09-704-640-71					
; Sequence 71, Application US/09704640					
; Patent No. 6635459					
GENERAL INFORMATION:					
; APPLICANT: Lewis, No. 6635459man G.					
; APPLICANT: Dinkova-Kostova, Albena T.					
; APPLICANT: Fujita, Masayuki					
; APPLICANT: Gang, David R.					
; APPLICANT: Sarkanen, Simo					
; APPLICANT: Ford, Joshua D					
TITLE OF INVENTION: RECOMBINANT PINOESINOL/LARICINESINOL REDUCTASES,					
FILE REFERENCE: WSUR-1-16492					
CURRENT APPLICATION NUMBER: US/09/704,640					
CURRENT FILING DATE: 2000-11-02					
PRIOR APPLICATION NUMBER: 09/475,316					
PRIOR FILING DATE: 1999-12-30					
PRIOR APPLICATION NUMBER: 09/307,653					
PRIOR FILING DATE: 1999-05-07					
PRIOR APPLICATION NUMBER: PCT/US97/20391					
PRIOR FILING DATE: 1997-11-07					
PRIOR APPLICATION NUMBER: 60/054,380					
PRIOR FILING DATE: 1997-07-31					
PRIOR APPLICATION NUMBER: 60/030,522					
PRIOR FILING DATE: 1996-11-08					
NUMBER OF SEQ ID NOS: 122					
SOFTWARE: PatentIn Ver. 2.0					
SEQ ID NO 71					
; LENGTH: 1328					
; TYPE: DNA					
; ORGANISM: Tsuga heterophylla					
; FEATURE:					
; NAME/KEY: CDS					
; LOCATION: (20)..(946)					
US-09-475-316A-71					
Alignment Scores:					
Pred. No.:	294	Length:	1328		
Score:	7.00	Matches:	7		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	4.93%	Indels:	0		
DB:	3				

EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,893  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,900  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,916  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,373  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 212  
LENGTH: 1551  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE

LOCATION: (420)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (1017)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (1408)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (1423)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-205-258-212  
Alignment Scores:  
Pred. No.: 339 Length: 1551  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.93% Indels: 0  
DB: 4 Gaps: 0  
US-09-690-825-34 (1-142) x US-09-205-258-212 (1-1551)  
QY 13 PheLeuLysAspHisArgile 19  
DB 177 TTTTGAAGATCATAGGATA 197  
RESULT 73  
US-09-648-004-9  
Sequence 9, Application US/09648004  
Patent No. 6498242  
GENERAL INFORMATION:  
APPLICANT: CHEN QIONG  
APPLICANT: THOMAS, STUART  
APPLICANT: NAGARAJAN, VASANTHA  
TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIC ACID AND  
TITLE OF INVENTION: INTERMEDIATES  
FILE REFERENCE: CL-1341-A  
CURRENT APPLICATION NUMBER: US/09/648,004  
CURRENT FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/252,553  
PRIOR FILING DATE: 1999-02-19  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 9  
LENGTH: 1644  
TYPE: DNA  
ORGANISM: Acinetobacter sp.  
US-09-648-004-9  
Alignment Scores:  
Pred. No.: 358 Length: 1644  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.93% Indels: 0  
DB: 4 Gaps: 0  
US-09-690-825-34 (1-142) x US-09-648-004-9 (1-1644)  
QY 129 LysLysValArgAlaAla 135  
DB 1226 AAGAAGGTCCTCGAGCTATA 1246  
RESULT 74  
US-09-107-532A-2080  
Sequence 2080, Application US/09107532A  
Patent No. 6563275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush

Tue Aug 17 05:54:51 2004

us-09-690-825-34.oligo.rni

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneks

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 2080:

SEQUENCE CHARACTERISTICS:

LENGTH: 1743 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...1743

SEQUENCE DESCRIPTION: SEQ ID NO: 2080:

US-09-107-532A-2080

Alignment Scores: 378 Length: 1743  
Pred. No.: 7.00 Matches: 7  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 4.93% Gaps: 0  
DB: 4

US-09-690-825-34 (1-142) x US-09-107-532A-2080 (1-1743)

Qy 111 AsnLysIleAlaLysGluThr 117

Db 679 AATAAAATAGCTAAGGAACG 699

RESULT 75

US-08-208-036-16

Sequence 16, Application US/08208036

Patent No. 5436326

GENERAL INFORMATION:

APPLICANT: Yoshizumi IISHINO et al.

TITLE OF INVENTION: METHOD FOR CLONING OF A GENE FOR POL I TYPE

TITLE OF INVENTION: DNA POLYMERASE

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack

STREET: 805 Fifteenth Street, N.W., #700

CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/208,036  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/897,282  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 1779 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: genomic DNA

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM: Bacillus caldotenax

STRAIN: YT-G(DSM406)

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE:

CELL TYPE:

CELL LINE:

ORGANELLE:

IMMEDIATE SOURCE:

LIBRARY:

CLONE:

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATE:

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO:

US-08-208-036-16

Alignment Scores: 385 Length: 1779  
Pred. No.: 7.00 Matches: 7  
Score:

Percent Similarity: 100.00%      Conservative: 0  
 Best Local Similarity: 100.00%      Mismatches: 0  
 Query Match: 4.93%      Indels: 0  
 DB: 1      Gaps: 0

US-09-690-825-34 (1-142) x US-08-208-036-16 (1-1779)

Qy 5 ThrLeuProAlaTrpGln 11

Db 876 ACATTACGCCGACGCTTGGCAA 896

Search completed: August 16, 2004, 03:24:53  
 Job time : 112 secs